UC Santa Cruz

UC Santa Cruz Electronic Theses and Dissertations

Title

When Human Universal Meets Language Specific

Permalink

https://escholarship.org/uc/item/00r1t0bb

Author

Guevara, Jed Sam Pizarro

Publication Date

2020

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

WHEN HUMAN UNIVERSAL MEETS LANGUAGE SPECIFIC

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

LINGUISTICS

by

Jed Sam Pizarro Guevara

June 2020

The Dissertation of Jed Sam Pizarro Guevara
is approved:

_____
Associate Professor Matthew Wagers, Chair

_____
Distinguished Professor Emerita Sandra Chung

_____
Assistant Professor Amanda Rysling

_____
Quentin Williams
Acting Vice Provost and Dean of Graduate Studies

Copyright © by

Jed Sam Pizarro Guevara

2020

Table of Contents

List of Figures vii

List of Tables viii

Abstract x

Dedication xii

Acknowledgments xiii

1 Introduction 1
- 1.1 The goals of the dissertation . . . 3
- 1.2 Outline of the dissertation . . . 6

2 Tagalog essentials 9
- 2.1 Language background . . . 10
- 2.2 Basic clause structure . . . 10
- 2.3 Morphological case and pronouns . . . 12
- 2.4 An excursus: Subjecthood in Tagalog . . . 14
- 2.5 Voice morphology . . . 16
- 2.6 The interaction of voice and word order . . . 17
- 2.7 Three syntactic constructions in Tagalog . . . 20
  - 2.7.1 *Ay*-inversion . . . 20
  - 2.7.2 Relative clauses . . . 22
  - 2.7.3 *Wh*-questions . . . 25
- 2.8 The interaction of voice and extraction . . . 28
- 2.9 The theoretical landscape . . . 30
  - 2.9.1 Approach 1: Multiple specifiers . . . 30
  - 2.9.2 Approach 2: Agreement with C . . . 35
- 2.10 Conclusion . . . 39

3 The extraction restriction, revisited **41**
3.1 The alternative: Voice-mismatch restriction . . . 42
3.2 Mismatches under non-AV voices "in the wild" . . . 47
3.3 Experiment 1: Comparing voice (mis)match extractions under AV and PV 51
3.3.1 Participants . . . 51
3.3.2 Materials . . . 52
3.3.3 Procedure . . . 52
3.3.4 Data analysis . . . 53
3.3.5 Results . . . 54
3.3.6 Discussion . . . 56
3.4 Experiments 2: Comparing voice (mis)match extractions under AV and PV again . . . 60
3.4.1 Participants and procedure . . . 60
3.4.2 Materials . . . 61
3.4.3 Data analysis . . . 64
3.4.4 Results . . . 64
3.4.5 Discussion . . . 66
3.5 Experiments 3: Comparing voice (mis)match extractions under applicatives 68
3.5.1 Participants and procedure . . . 68
3.5.2 Materials . . . 68
3.5.3 Data analysis . . . 69
3.5.4 Results . . . 72
3.5.5 Discussion . . . 74
3.6 A model of the participants' judgment process . . . 75
3.6.1 Method . . . 76
3.6.2 Judgment process in Experiment 1 . . . 78
3.6.3 Judgment process in Experiment 2 . . . 79
3.6.4 Judgment process in Experiment 3 . . . 80
3.6.5 Discussion . . . 82
3.7 General discussion and conclusion . . . 82
3.7.1 A reformulation: Two grammars? . . . 83
3.7.2 Implications for formal analyses of the extraction restriction . . . 85
3.7.3 Implications for sentence processing . . . 100

4 The role of voice in dependency formation **103**
4.1 Comprehenders actively construct dependencies . . . 105
4.2 A hypothesis . . . 108
4.3 Experiment 4: Isolating the contribution of voice . . . 113
4.3.1 Voiceless verbs impose similar restrictions on movement . . . 114
4.3.2 Participants . . . 116
4.3.3 Materials . . . 116
4.3.4 Procedure . . . 119

4.3.5 Analysis . . . 119
4.3.6 Results . . . 121
4.3.7 Discussion . . . 123
4.4 Experiment 5: Comparing voice cues . . . 124
4.4.1 Participants . . . 125
4.4.2 Materials, procedure, and analysis . . . 125
4.4.3 Results and discussion . . . 126
4.4.4 Discussion . . . 132
4.5 General discussion . . . 133
4.5.1 Tagalog comprehenders attend to voice information at the FGD-level . . . 134
4.5.2 Unrejected reading times: A more nuanced picture of choice . . . 137
4.5.3 (A)symmetries as a window to interactions . . . 142
4.5.4 Issues and future directions . . . 144

5 The role of word order in relative clause processing 148
5.1 Subject-object asymmetry in relative clauses . . . 150
5.2 Accounting for the asymmetry . . . 151
5.2.1 The role of intervening material . . . 152
5.2.2 The role of linguistic experience . . . 156
5.3 Attenuating the asymmetry . . . 158
5.3.1 Animacy of the head noun . . . 158
5.3.2 Referential type of intervening elements . . . 159
5.3.3 Relative order of the head noun and relative clause . . . 161
5.4 The present studies: An overview . . . 166
5.4.1 Participants . . . 167
5.4.2 Procedure . . . 167
5.4.3 Data analysis . . . 169
5.4.4 Preview of the results . . . 170
5.5 Experiment 6: Unambiguous RCs with R-expression co-arguments . . . 171
5.5.1 Materials . . . 171
5.5.2 Predictions . . . 172
5.5.3 Data analysis . . . 176
5.5.4 Results and discussion . . . 176
5.6 Experiment 7: Unambiguous RCs with pronominal co-arguments . . . 180
5.6.1 Materials . . . 181
5.6.2 Prediction . . . 181
5.6.3 Data analysis . . . 186
5.6.4 Results and discussion . . . 187
5.7 Experiment 8: Ambiguous RCs with R-expression co-arguments . . . 190
5.7.1 Materials . . . 192
5.7.2 Prediction . . . 192

5.7.3 Data analysis 194
5.7.4 Results and discussion 195
5.8 General discussion 199

6 Conclusion **203**
6.1 The extraction restriction is not clean-cut 204
6.2 Voice is a cue used in FGD-processing 205
6.3 SRCs are easier to process than ORCs in Tagalog 206
6.4 Future directions and concluding remarks 207
6.4.1 Two grammars of Tagalog?: Collecting child data 207
6.4.2 Do different FGDs exhibit varying mismatch penalties? 209
6.4.3 Voice mismatches in islands 211
6.4.4 What other cues do Tagalog comprehenders use? 213
6.4.5 Quantifying how informative voice is 215
6.4.6 Eye-movement in relative clause processing 216

References **219**

List of Figures

2.1 Map of the Tagalog-speaking regions of the Philippines. 11

3.1 Mean ratings by VOICE, MATCH, and TYPE in Experiment 1 55
3.2 Mean ratings by VOICE, MATCH, and TYPE in Experiment 2 65
3.3 Mean ratings by AGEX, MATCH, and TYPE in Experiment 3 72
3.4 Two potential underlying distributions of intermediate ratings 76

4.1 A schema of how a comprehender could use voice to guide interpretation. 110
4.2 Phrase-by-phrase discrimination scores in experiment 4 122
4.3 Phrase-by-phrase discrimination scores in Experiment 5 130

5.1 A schematization of a typical trial in Experiments 6–8 168
5.2 Breakdown of responses by accuracy and confidence in Experiment 6 . . 177
5.3 RTs of correct responses in Experiment 6 178
5.4 Breakdown of responses by accuracy and confidence in Experiment 7 . . 187
5.5 RTs of correct responses in Experiment 7 188
5.6 Breakdown of responses by accuracy and confidence in Experiment 8 . . 196

6.1 Sample item of the proposed reward/punishment judgment study for children 210
6.2 Sample picture selection screen used in the experiments presented in Chapter 5 217

List of Tables

2.1 Case marker allomorphy 13
2.2 Pronominal allomorphy 14
2.3 The bifurcation of subject-properties in Tagalog 15

3.1 Schematization of the items used in Experiment 1 53
3.2 Descriptive statistics of Experiment 1 54
3.3 Summary of cumulative link mixed models in Experiment 1 56
3.4 Sample experimental items used in Experiment 2 62
3.6 Descriptive statistics of Experiment 2 64
3.7 Summary of the cumulative link mixed models estimated in Experiment 2 66
3.8 Mapping between conditions and extraction in Experiment 3 69
3.9 Sample experimental items used in Experiment 3 70
3.11 Descriptive statistics of Experiment 3 72
3.12 Summary of the cumulative link mixed models estimated in Experiment 3 73
3.13 Descriptive statistics of filler items in Experiments 2 and 3 77
3.14 Summary of the response models for Experiment 1 79
3.15 Summary of the response models for Experiment 2 80
3.16 Summary of the response models for Experiment 3 81
3.17 A summary of when C probes together or separately in Tagalog under a ϕ-agreement approach to voice and the extraction restriction 99

4.1 Sample item used in Experiment 4 117
4.2 Sample experimental items used in experiment 5 127

4.5 Discrimination behavior and different levels of granularity 135
4.6 Descriptive statistics of unrejected reading times in Experiment 5 139
4.7 Summary of linear mixed-effects models estimated in Experiment 5 . . . 140

5.1 Frequencies of each type of relative clauses in English 157
5.2 Summary of excluded data (in %) by exclusion-type for Experiments 6–9 . 170
5.3 Preview of the results (Experiments 6–9) 171
5.4 Sample experimental items used in Experiment 6 172
5.5 Comparison of the relative order of S and O in main clauses and relative clauses in Tagalog 175
5.6 Summary of cumulative link mixed models in Experiment 6 178
5.7 Summary of linear mixed-effects models in Experiment 6 179
5.8 Summary of Experiment 6: How does each account fare? 180
5.9 Sample experimental items used in Experiment 7 182
5.10 Summary of cumulative link mixed models in Experiment 7 188
5.11 Summary of linear mixed-effects models in Experiment 7 189
5.12 Summary of Experiment 7: How does each account fare? 190
5.13 Sample experimental items used in Experiment 8 193
5.14 Summary of cumulative link mixed models in Experiment 8 196
5.15 Summary of linear mixed-effects models in Experiment 8 198
5.16 Summary of Experiment 8: How does each account fare? 199

6.1 Sample experimental items used to isolate the underlying cause for the inconsistencies 212

Abstract

When human universal meets language specific

by

Jed Sam Pizarro Guevara

In the broadest sense, this dissertation is an effort to better understand human linguistic capacity. Empirically, I focus on Tagalog filler-gap dependencies (FGDs) and how comprehenders understand them. To successfully interpret FGDs, comprehenders must encode in memory the elements that form the dependency, maintain the grammatical and thematic relations between these elements, and the licensing requirements of the elements encountered thus far throughout the course of interpretation. These are *some* of what we do when comprehending language more generally. What appears to be an effortless process is in fact a complex ability involving interactions between different classes of information—some linguistic, some, not.

The primary goal of this dissertation is to better understand the classes of linguistic information that comprehenders attend to and use when processing FGDs. I examine this question in two contexts: (i) the role of morphological information like voice morphology in the incremental processing of Tagalog FGDs; and (ii) the role of syntactic information like word order, as well as the role of intervening linguistic material and linguistic experience, in relative clause processing. The secondary goal of this dissertation is to contribute to the growing efforts to incorporate data from "smaller" languages in psycholinguistic research to directly address the lack of linguistic diversity that the field suffers from.

The main insight of this dissertation is that Tagalog, a language that is typologically different from the more familiar languages well-characterized in psycholinguistics, exhibits what seems to be universal properties of language comprehension: (i) comprehenders actively associate a filler to a gap even before the availability of disambiguating information; and (ii) they prefer relative clauses with subject gaps over relative clauses with

object gaps. There are, however, language specific properties that can modulate how actively comprehenders link the filler to the gap—like finer grained information about the dependencies, for example—or that can attenuate the robust preference for subject gaps in relative clauses—like the relative order of the head noun and the relative clause, for instance.

Para sa akong inahan, amahan, ug igsuon

Acknowledgments

While this dissertation is under my name, this in reality is a product of the many interactions I have had with various people throughout my academic career. I apologize in advance because this will be long. I just have so many people to thank. Sorry, not sorry.

I cannot thank Matt Wagers enough for his guidance, patience, and generosity—not just with his time, but also with his resources—throughout the years. He has taught me almost everything I know about psycholinguistics. He has given me countless opportunities to learn how to be a field psycholinguist through first-hand experience. Because of his unwavering support both intellectually and financially, I was able to go the Philippines for research almost every summer since I started the program. Most, if not all, of the projects I have done in graduate school—even the ones that were supposed to be outside his wheelhouse—have benefited from discussions with him. If/when I ever have advisees of my own some day, I would support them MWW-style. If this dissertation is supposed to show how grateful I am to Matt, then I'm afraid I have fallen short twice over.

I also want to thank Sandy Chung for showing me what responsible fieldwork looks like. I truly admire her careful description and in-depth investigation of Chamorro, along with her sustained involvement in the community. When I think about the type of fieldworker that I want to be both now and in the future, I try to emulate her. One of the lessons that I learned from her very early on is that it's okay to "adopt something with a smile." I also want to thank her for collaborating with Matt and producing that *Language* paper that started it all for me, that made me want to become a field psycholinguist.

Even though Amanda Rysling entered my academic life rather late in the game, her contributions to my growth have been nothing short of profound. To use words from Miley Cyrus' greatest hit, Amanda "came in like a wrecking ball." She has influenced how I design my experiments and she's definitely made me more paranoid—in a good way. She's taught me not to assume that every problem is a nail, just because I've only ever

used hammers. She's made me consider and really take seriously low-level task effects and the role of decision making in language comprehension. She was also one of the emotional doulas I had during the job application process, giving me feedback on my materials while listening to me bitch and moan about academia. I value the conversations I've had with her—and her Din Tai Fung recommendations, too.

I also want to thank the other faculty members for creating this tightly knit Linguistics community that I have had a pleasure of being a part of. Jim McCloksey deserves a special mention. While he was never really an advisor in any official capacity, he was very supportive and nurturing. He would take a very modest observation I've made about whatever topic and somehow turn that observation into gold. From him I learned how to be charitable—to give any proposal its best chance to succeed. My only regret is that I never got around to working more on Dabaw Bisaya's verum focus other than that very preliminary investigation I did for his class on polarity.

Many thanks to Jake Vincent, who is probably one of the closest friends I have made in graduate school. We're so different, yet so alike. He's quiet; I am quite boisterous. We both care so much about the overall aesthetic of documents. Everything has to be calculated down to the last pixel. He made teaching Ling 154 bearable. He was an excellent TA, even though I never got around to filling out his TA eval. Oops. I'm grateful for his being patient with me as I fumbled around trying to figure out how to teach that class. We've also been co-TAs for two classes and his level of organization is beyond disgusting—and I mean this in a very positive way. I applaud his resilience and his ability to take the full brunt of my bitching and moaning about academia, my insecurities and incompetence, and other self-loathing soliloquies.

I was (un)fortunate enough to be a part of a really capable cohort—and I would even say the best dressed—and that often made me question whether I should even be in graduate school. With the help of the same people whose talents I was really intimidated by

(and other people, of course), I'm here now writing the acknowledgment section of my dissertation. I want to thank Margaret Kroll and Steven Foley for the friendship all these years, and for the bi-weekly drinks during job application season. It made all the rejections a lot more bearable when I had my friends to commiserate with me. I also want to thank my cohort mates who left earlier than us: Hitomi Hirayama Tomida, Nate Clair, Ben Mericli, and Jeff Adler for their friendship early on in the program. Chelsea Miller deserves a shout out. My productivity was quite high early on in the program because I fed off her anxious energy. I want to thank her for letting me stay with her when I was looking for housing, for resurrecting s/lab with me, and for being a fabulous friend all these years. If she can see how much s/lab has grown, she'd be a proud mama.

Many thanks to my fellow residents of Bethany's Curves. Kelsey Sasaki held the Curve together. She was definitely the parent of the Curve: Mamasaki 2.0. She's also an excellent office mate, always trying to fatten me up with her chocolates and sugary goodness! Her Gudetama stress ball(?) is so fun to poke and squeeze and squeeze and poked it I have—many times, actually, while writing this dissertation. Nick Van Handel listened to many of my self-loathing soliloquies, read excerpts of what I've written, spilled tea and threw shade with me, and perhaps most importantly, was almost always down to eat—especially when it was Indian buffet. I associate him with food coma: after Mumbai or Ambrosia, we would, without fail, end up taking the entire afternoon off, doing absolutely nothing other than recovering from food coma even if we had initially agreed to do work after.

I also thank Netta Ben-Meir Sokolov for the friendship all these years (and for the really cute pics of baby Roro that always made me smile). Like Nick, she's also listened to my self-loathing soliloquies, read excerpts of what I've written, and spilled tea and threw shade with me. I also enjoy talking about food and TV shows with her. Thanks for the recommendations. I also thank Stephanie Rich. We haven't been friends for that long, but it feels like a very long time. I want to thank her for keeping me company (virtually, and

in person with good social distancing, of course) during this pandemic. When Chelsea left, I thought there wasn't going to be a ball of anxious energy to feed off of. Then she decided to come here!

Many thanks to the other grad students who've made this journey fun and exciting (in chronological order): to Deniz Rudin for always being down to drive anywhere for good food; to Kelsey Kraus for her services to the department even after she's graduated; to Jenny Bellik for showing me what work-life balance is supposed to look like and for her FB posts detailing her children's language acquisition; to Maho Morimoto for the food- and work-dates; to Tom Roberts for his love for boardgames and JRPGs; to Anissa Zaitsu for being the best of both worlds (couldn't really think of a Lizzie McGuire song, so here's Hannah Montana), showing me that smart people know how to have fun, too, and that fun people can be smart, too; to Ben Eischens, "the muscle TA," for practically having the same upbringing as me (LOL) and for checking in on me constantly during my last sprint to finish this dissertation; to Jérémie Beauchamp for being a pleasant co-TA in Semantics and for almost always being down to nom noms wherever; to Richard Bibbs for taking time out of his busy schedule to be our DM; to Jack Duff for coordinating s/lab and for the many laughs during DnD; Max Kaplan for his font recommendations and for his lovely family, Kara and Cricket; to Dan Brodkin for his interests in anything Austronesian and reviving AnG (now NAnG); to Vishal Arvindam for the recommendations on how to analyze Visual World data; and to Lalitha Balachandran for working-while-drinking with me. Many thanks to the other graduate students not mentioned here for being a part of this long and arduous journey.

Many of us grad students take for granted how functional our department is. We don't realize this until we hear the horror stories from other departments at UCSC or other Linguistics departments. We often forget that we owe how functional our department is to our staff. Much gratitude to Ashley Hardisty, Maria Zimmer, Logan Roberts, and

Gwyn Vandevere—and to Matt MacLeod, too, when he still worked here.

I also thank my undergrad research assistants, Miguel Toner and Sikander Haider, for helping me design and make experimental items and for running the locative inversion experiments. Even though these experiments did not make it into this dissertation unfortunately, I want to thank them for being patient with me as I fumbled around, trying to figure out how to be a mentor. It was a journey, but I think we got there. My only hope is that that entire experience was as rewarding for them as it was for me.

I have become friends with many of my undergrad students, but many thanks to two very talented women for their friendship and conversations, Sonia Dominguez and Claire Miller. I truly enjoy our chats after Ling 154 and Ling 180, when I was not their instructor but just another person in the department. I value their drive to succeed and their passion to diversify academia. Many thanks to Gabi Pleasant for her infectious bubbly personality, to Jackson Confer for random conversations and perhaps more importantly, conversations about Civ 6, and to Spencer Gilbert for always being down to grab Pericos.

I want to thank my Asian-American Pacific Islander Resource Center (AAPIRC) family for giving me a sense of community (outside of Linguistics) when I needed it the most. I am grateful for Nancy Kim, the director of the resource center then, for providing this resource to our community and for giving me an opportunity to grow as a person. I learned a lot in my one year at AAPIRC: about myself, about our community's history, and how to work with others. It's okay to not control everything. It was not an easy journey to relinquish control, but I think I got there. My time at AAPIRC definitely helped me form meaningful relationships with my RAs through mentorship. I also thank Jacob Velasquez for supporting Nancy and us interns; Andi Min, the intern that I'm closest with, for the laughs, tears, and conversations; and of course, the rest of the interns who I have had the pleasure of working with. Special mention to Collette Quach, Srija Srinivasan, Cami Shigaya, and Brian Sou.

I want to thank Rosa Baeza and Brian Oliva for 15+ years of friendship—and counting. Sometimes, we go for long periods of time without talking because we're all so busy with our lives. I know that as soon as we do start talking, it's as if no time has passed and we can just pick up where we left off. I know that they'll always have my back, and I have theirs. I thank them for their love and support throughout the years.

My gratitude extends to Alice Shen, Raksit Lau, and Adam Singerman for being my remote job-applications-and-dissertation-writing doulas. Their friendship, despite the distance (okay, not really true for the first two since they're only an hour and a half drive) has made this getting-kicked-out-of-the-nest-experience bearable. I want to thank them for listening to me bitch and moan about my insecurities, incompetence, and other self-loathing soliloquies. Hmmmm. These seem to be a recurring theme. Ha!

I am indebted to many Filipino academics from other institutions: to Soleil Davíd for being my friend since college and for answering very obscure questions about Tagalog; to Henrison Hsieh for conversations about Tagalog morphosyntax, both the relevant and not-so-relevant; and to Nozomi Tanaka and Rowena "Weng" Garcia for the pep talks, and the conversations about sentence processing. I've really only met Henrison, Nozomi, and Weng once or twice in real life but I feel like we've been friends for a very long time.

I am of course indebted to the Department of Linguistics at the University of the Philippines – Diliman. All the experiments reported here were collected while I was a Research Fellow for the Department. I want to thank the Department for giving me the opportunity to be a part of such a wonderful community; for helping me with data collection; and for being my home away from home. Special mention to Farah Cunanan and Madel Landicho for showing me the beauty of getting a massage after a stressful week and for all of the *chismis*; to Vicky Vidal for her motherly compassion towards anyone affiliated with the Department; to Aldrin Lee for allowing me to be a part of the UPLingg community intellectually; to Michael Manahan for playing AOV with me; to Divine En-

driga and Mantha Sadural for helping me compile the experiments presented in Chapter 5 and for the food dates; and of course, to Kristina Gallego for the many conversations about grad school, for the countless judgments I have asked of her to make my experimental items, and for her friendship all these years. *Maraming salamat, UPLingg. Makikigulo ulit ako in 2 years or so. Maglelechon tayo! Ako ang taya.*

Lastly, I am grateful for my family, who even though I have chosen a path definitely not as lucrative as one in the medical field as they had initially envisioned, still urged me to follow my heart and to pursue something I liked. *Charot! Maka-noseebleed man uy!* We might not talk much, or see eye to eye on so many things, but their love and support made this possible. *Mama, Papa, ug Franco, para sa inyo ni. Daghang salamat sa tanan.*

And now for the part that the thesis office wants dissertators to include in the acknowledgment section. I don't really understand why, but here it is. The text of this dissertation includes reprints of previously published material. Experiment 1, presented in Chapter 3, and Experiments 4 and 5, presented in Chapter 4, have been packaged together and published in Frontiers in Psychology:

Pizarro-Guevara, J.S. & Wagers, M. (2020). The predictive value of Tagalog voice morphology in filler-gap dependency formation. *Frontiers in Psychology, 11*, 517. `https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00517`

The co-author listed in this publication directed and supervised the research which forms the basis for part of this dissertation.

The rest of Chapter 3 and all of Chapter 5 are also being prepared for publications:

Pizarro-Guevara, J.S. & Wagers, M. *A tale of two grammars of extraction in Tagalog: Evidence from acceptability judgment models*. Manuscript in preparation.

Pizarro-Guevara, J.S. & Wagers, M. *(A)symmetries in relative clause processing: A case study of pre- and post-nominal relative clauses in Tagalog*. Manuscript in preparation.

Chapter 1

Introduction

Languages all over the world exhibit a wide range of long-distance dependencies. Consider the sentences in (1). One example of a long-distance dependency is the relationship between a pronoun and its antecedent, as in (1-a). As English speakers, we know that the pronouns, *he* and *his*, may—though they need not—refer to the same entity, namely *Mike*, even though they surface quite far from one another. Another example is the relationship between a *wh*-word like *what* and the place where it is interpreted, as in (1-b). Again, as English speakers, we can intuit that the word *what* in the beginning of the sentence refers to the direct object of the verb *ate* (i.e., the thing that Pedro ate) even though *what* appears well to the left of *ate*.

(1) a. Mike was very happy when he told Lee that Peggy Sue is marrying his brother.

b. What did Maria say that Pedro ate last night?

The empirical focus of this dissertation is a type of long-distance dependency called FILLER-GAP DEPENDENCIES (FGDs). *Wh*-questions, as in (1-b), are an example. FGDs are syntactic constructions where a constituent appears not in its base-generated position (i.e., where it is interpreted), but rather to the left of the verb it is associated with. Consider (2-a): *the*

defendant's assets is in its base-generated position, after the verb *confiscated* where direct objects generally appear. Now contrast this with (2-b): no nominal appears after *confiscated* even though one would normally expect to find one in this position. Instead, the corresponding nominal has been displaced, and now occur before the verb. Throughout, I refer to the displaced nominal as the FILLER.

(2) a. The journalists said that the government confiscated the defendant's assets illegally.

b. The journalists wrote about the defendant's assets that the government confiscated ___ illegally.

Even though the filler has been displaced, we as comprehenders still interpret it as the direct object of the verb. That is, we seem to understand the filler to correspond to an empty position after the verb. Throughout, I refer to this empty position as the GAP, and represent it as "___" above. The relationship between the filler and the gap is what I refer to as a FILLER-GAP DEPENDENCY (FGD).

To successfully interpret FGDs, we as comprehenders must encode in memory the elements that form the dependency, maintain the grammatical and thematic relations (i.e., who is doing what to whom) between these elements, the licensing requirements of the elements encountered thus far throughout the course of interpreting a sentence. These are *some* of what we do when comprehending language more generally. What appears to be an effortless process is in fact a complex ability involving interactions between different classes of information—some linguistic, some, not—and different cognitive processes.

In the broadest sense, this dissertation is an effort to better understand human linguistic capacity. Particularly relevant to my dissertation is the claim that we share the same cognitive architecture (*e.g.,* the function of our working memory and attention), despite the superficial differences that exist in the world's languages. My dissertation is entitled

When human universal meets language specific because I am particularly interested in the question of how our cognitive architecture (the human universal) is affected by the superficial differences that exist in the world's languages (the language specifics). I examine this question in the context of how FGDs are comprehended in Tagalog, a Central Philippine language that belongs to the Malayo-Polynesian sub-group of the Austronesian language family (Eberhard, Simons, & Fennig, 2019) and that is spoken by approximately 22.5 million native speakers (Philippine Statistics Authority, 2017).

As we will see, Tagalog has an elaborate system of verbal morphology that interacts with case marking and is intimately correlated with what arguments can participate in FGDs. The language also exhibits a great deal of word order flexibility, not just at the clausal level (i.e., how co-arguments are ordered with respect to the verb), but also at a much smaller level (e.g., in the nominal domain). Particularly relevant here is the flexibility in relative clauses, in terms of where the head noun surfaces relative to the RC. How do these morphosyntactic features affect how Tagalog comprehenders understand FGDs?

1.1 The goals of the dissertation

The dissertation has two main goals. The first is to better understand what classes of (linguistic) information we as comprehenders attend to and use when processing filler-gap dependencies. In this dissertation, I examine this question in two contexts:

- in chapter 4 where I investigate the contribution of morphological information like voice morphology in Tagalog on incremental parsing;
- in chapter 5 where I investigate the role of syntactic information like word order, as well as the role of intervening linguistic material and linguistic experience, in relative clause processing.

It is *always* an empirical question what classes of linguistic information we as com-

prehenders attend to and actually use as cues, and how we use them in real-time sentence processing. We know from the grammatical illusion literature—like agreement attraction (Clifton, Frazier, & Deevy, 1999; Pearlmutter, Garnsey, & Bock, 1999; Wagers, Lau, & Phillips, 2009) and illusory licensing of negative polarity items (Drenhaus, Saddy, & Frisch, 2005; Vasishth, Brüssow, Lewis, & Drenhaus, 2008; Xiang, Dillon, & Phillips, 2006), to name a few—that comprehenders are sometimes sensitive to irrelevant cues. By irrelevant, I mean that these are cues that are not instantiated by the dependency of interest. In the case of agreement attraction, the irrelevant cue is the number feature of an intervening noun that is part of the subject noun phrase. In the case of illusory licensing of negative polarity items, the irrelevant cue is a non-c-commanding negation. Furthermore, we also know from the acquisition literature that comprehenders can also ignore very reliable cues and at times, even do so in favor of less reliable ones (Gagliardi & Lidz, 2014; MacWhinney, Bates, & Kleigl, 1984). Thus, even if some linguistic feature of a given language is a cue that is available in the input and is generally reliable in computing the interpretation of a given dependency, whether comprehenders of a given language actually use it remains an open question. How they will use it is another open question. Simply put, cue availability and reliability do not necessarily entail use.

The second goal of this dissertation is to contribute to the growing efforts to incorporate data from "smaller" languages in psycholinguistic research. This enterprise has been referred to as field psycholinguistics (Christianson & Cho, 2009) or crosslinguistic psycholinguistics (Norcliffe, Harris, & Jaeger, 2015). Much of what we know about sentence processing suffers from a lack of linguistic diversity, since the data come primarily from Western (and some East Asian), educated, industrialized, rich, and democratic language users ("WEIRD"; Henrich, Heine, & Norenzayan, 2010). In fact, only 10 languages make up the bulk of the research—approximately 85%—on language processing (Anand, Chung, & Wagers, 2011). According to Ethnologue (Eberhard et al., 2019), the Austrone-

sian language family accounts for 5–17% of the world's linguistic diversity, depending on whether we quantify linguistic diversity based on the number of extant languages or on the number of speakers. It is curious then that this language family is "virtually absent from psycholinguistic research" (Wagers, Borja, & Chung, 2015). Thus, this dissertation is of theoretical import because it addresses directly this lacuna by investigating Tagalog, a verb-initial language spoken primarily in the Philippines by approximately 22.5 million native speakers (Philippine Statistics Authority, 2017). Tagalog is a language that is typologically different from the more familiar languages that are well-characterized in the psycholinguistics literature.

In recent years, there has been a slow but steady accumulation of psycholinguistic work focusing on typologically diverse languages. For a review, see Norcliffe et al. (2015). Relevant to the second goal of this dissertation and its focus on Tagalog are six studies on Tagalog sentence comprehension. Sauppe (2016) used the visual word paradigm to investigate whether Tagalog comprehenders anticipated the referents of upcoming post-verbal arguments based on syntactic function or semantic role in canonical declaratives (i.e., verb-initial sentences). He found that comprehenders leveraged the thematic relations of the verb to identify the referents and argued that event semantics takes temporal precedence during comprehension.

Meanwhile, Garcia and colleagues used self-paced listening with picture-selection (Garcia, Roeser, & Höhle, 2019) and eye-tracking while listening (Garcia, Roeser, & Höhle, 2020) to investigate whether comprehenders leveraged word order and morphosyntactic markers (i.e., case and voice) in the comprehension of canonical declaratives. They found that it was easier for children to understand agent-initial sentences compared to patient-initial sentences. This word order effect, however, was modulated by voice.

The experiments presented in Chapter 4 differ from these studies in two ways. First, the previous studies have visual cues available to the comprehender that could give them

an overview of the global event structure; mine did not. Second, my experimental items involve FGDs; theirs did not. The relative simplicity of their experimental items, coupled with a rich paralinguistic context, might reduce the need for comprehenders to deploy resources from other parts of the grammar. What Chapter 4 is interested in is whether voice will be deployed as a cue in contexts where an argument is displaced from its thematic position and has to be held in working memory. The experiments presented in Chapter 5 also have visual cues, but used items that involved FGDs.

In earlier work (Pizarro-Guevara, 2014), I investigated how relative clauses are processed using a referent-matching task, a modified version of a picture-matching task, modeled after Arnon (2010). More recently, Tanaka, O'Grady, Deen, and Bondoc (2019), Bondoc, O'Grady, Deen, and Tanaka (2018) also used a referent-matching task to investigate the comprehension of relative clauses in neurotypical and aphasic populations, respectively. While these studies do involve FGDs, they are different from the studies presented in this dissertation in two ways. First, the experiments presented in Chapter 4 are interested in the incremental choices comprehenders make as more of the input becomes available. The results of the previous studies are based on interpretations that are the end state of comprehension. Second, the experiments in Chapter 5 are in a sense a replication and an extension of these previous studies, which focused only on the comprehension of head-initial relative clauses.

1.2 Outline of the dissertation

The dissertation is organized as follows:

In Chapter 2, I provide an overview of the morphosyntactic features of Tagalog that figure prominently in this dissertation. More specifically, I present the received empirical generalization about Tagalog voice morphology and its interaction with movement dependencies: only arguments cross-referenced by voice are eligible for extraction. I

also survey four recent formal approaches to voice morphology and the received view of the extraction restriction (Aldridge, 2008; Rackowski & Richards, 2005; Aldridge, 2018; V. Chen, 2017).

In Chapter 3, I revisit the interaction between voice morphology and movement dependencies, and I question the robustness of the received generalization. It has been observed that some speakers productively allow the extraction of the agent even when this argument is not cross-referenced by voice. These "exceptions" have been documented by Philippine linguists before (Ceña & Nolasco, 2011, 2012). However, it has gone largely undiscussed in mainstream linguistics since these works were written in Tagalog. I make these descriptions available for wider consumption, alongside data from my own fieldwork and naturally occurring data I mined from various sources online, ranging from social media posts to online newspapers. I also present a series of acceptability judgment studies, which are, to my knowledge, the first large-scale investigation on this topic, and computational evidence using the framework developed by Dillon, Staub, Levy, and Clifton (2017). Results from these judgment studies are consistent with a view of the extraction restriction that is not as clean-cut. I suggest that the competing generalizations can be reframed terms of grammar competition à la Han, Lidz, and Musolino (2007).

In Chapter 4, I examine the functional contribution of voice morphology in real-time FGD-processing using the Stops-Making-Sense paradigm (Boland, Tanenhaus, & Garnsey, 1990; Boland, Tanenhaus, Garnsey, & Carlson, 1995) to detect plausibility mismatch effects, which in turn, could give us a window into the time course of dependency formation. I first compare the time course of FGD-formation when the verb exhibits voice morphology and when it does not in order to isolate the contribution of voice morphology in FGD-processing. Then I compare the time course of FGD-formation when the verb exhibits different flavors of voice to see the different ways that voice can be leveraged in FGD-processing. I add Tagalog to the list of languages supporting active dependency for-

mation. Comprehenders actively link the filler to the gap and they can use morphological information, like voice morphology in Tagalog, to do so.

In Chapter 5, I leverage the fact that Tagalog has both unambiguous and ambiguous RCs to investigate how RCs with subject gaps and RCs with object gaps are processed in the language. This chapter is in a sense a replication and an extension of the previous RC studies in Tagalog (Pizarro-Guevara, 2014; Bondoc et al., 2018; Tanaka et al., 2019). I also leverage the fact that Tagalog allows head-initial and head-final RCs to investigate whether and how word order can affect how RCs are processed. I add Tagalog to the list of languages exhibiting a subject-object asymmetry in RC processing. Comprehenders prefer RCs with subject gaps over RCs with object gaps. I also add to the body of literature that investigates what factors can modulate the asymmetry. I present a language-internal comparison à la Wagers, Borja, and Chung (2018) that suggests that in Tagalog, the relative order of the head noun and the RC can attenuate the asymmetry.

Finally, in Chapter 6, I conclude.

Chapter 2

Tagalog essentials

The main goal of this chapter is to describe the morphosyntactic features of Tagalog that will figure in the experiments in Chapters 3 to 5. In a sense, this chapter is a compilation of what we know about Tagalog from earlier descriptions of the language, insights from generative linguistics, and my own work with speakers from Metro Manila.

This chapter is organized as follows. I first provide a background of the language in Section 2.1, then an overview of the basic clause structure in Section 2.2 and morphological case in Section 2.3. I briefly digress in Section 2.4 to outline the problem of subjecthood in the Philippine context and make explicit what I mean by "subject" throughout this dissertation. In Section 2.5, I provide an overview of one of the central features of the language that is featured across different chapters, the language's voice morphology, and in Section 2.6, how it interacts with word order. In Section 2.7, I describe three syntactic constructions—*ay*-inverted sentences, relative clauses, and *wh*-questions—that are featured in the judgment studies in Chapter 3, the stops-making-sense experiments in Chapter 4, and the picture-selection experiments in Chapter 5. In Section 2.8, I discuss the received view of how voice morphology interacts with the movement dependencies surveyed previously. In Section 2.9, I review the four most recent approaches to voice morphology and its interaction with movement dependencies.

2.1 Language background

Tagalog is a Central Philippine language that belongs to the Malayo-Polynesian subgroup of the Austronesian language family (Eberhard et al., 2019). It is spoken by approximately 22.5 million native speakers (Philippine Statistics Authority, 2017)—most of whom are in the southern part of Luzon, one of the main islands of the Philippine archipelago. Figure 2.1 provides a map of the Philippines; shaded in blue are the regions in which Tagalog is spoken as the first language.[1]

This estimate of the number of Tagalog speakers increases dramatically if speakers of regional languages are included since most of them are early sequential bilinguals. They learn the language of the region or the province where they are from, and upon entry into the Philippine educational system, they learn Tagalog and English (Galang, 1988, 2001). Even though the Philippine Constitution of 1987 declares Filipino as the national language, the Tagalog/Filipino distinction is a non-issue for the purposes of this dissertation. I treat Tagalog and Filipino as one and the same. Filipino is simply the standardized form of Tagalog that allows borrowings from other regional languages (e.g., Cebuano) and foreign languages (e.g., English). For readers interested in other language ideologies surrounding the Tagalog/Filipino distinction, I recommend Appendix 2 of Barrios (2014), who provides a historical overview of how Filipino, the national language of the Philippines, came to be.

2.2 Basic clause structure

Tagalog is a head-initial language in which the predicate comes first in the clause. The language allows predicates of any category-type: verbal, as in (1-a); nominal, as in (1-b); adjectival, as in (1-c); or locative, as in (1-d).

[1]The map of Tagalog-speaking regions provided below is entitled "Idioma tágalo" and is by Fobos92. It is licensed under the Creative Commons Attribution-Share Alike 4.0 Unported license.

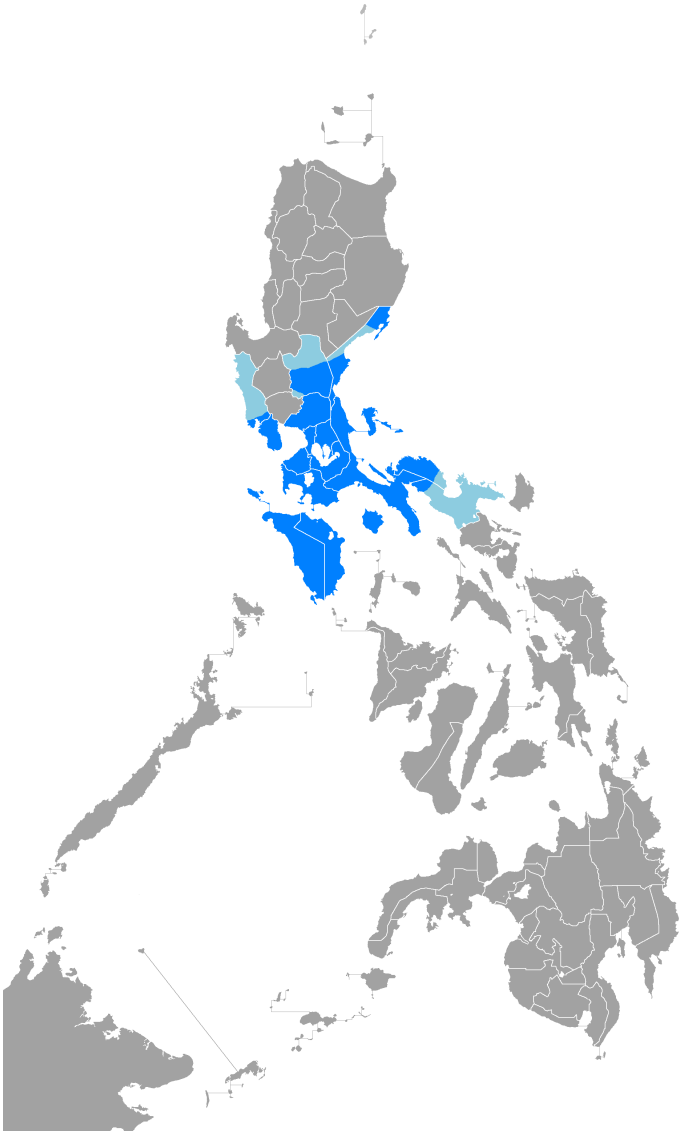


Figure 2.1: Map of the Tagalog-speaking regions of the Philippines.

(1) Adapted from Richards (2009)

a. Nagaaral ako
study.AV 1SG.NOM
'I am studying' *Verbal predicate*

b. Doktor ako
doctor 1SG.NOM
'I'm a doctor' *Nominal predicate*

c. Maganda ako
beautiful 1SG.NOM
'I am beautiful' *Adjectival predicate*

d. Nasa gitna ako
LOC middle 1SG.NOM
'I am in the middle' *Locative predicate*

In what follows, only verbal predicates are of interest. The order of post-verbal elements is relatively flexible. That is, they can scramble and still retain the grammaticality and the basic meaning of the sentence (Schachter & Otanes, 1983). This flexibility in post-verbal word order is exemplified in (2). In (2-a), the noun that functions as the direct object in the English translation is immediately after the verb. It is followed by the noun that functions as the indirect object, which in turn, is followed by the noun that functions as the subject. The other permutations of the arguments are shown in (2-b)–(2-f).

(2) Adapted from Schachter and Otanes (1983, p. 83)

a. Nagbigay ng libro sa babae ang lalaki
gave.AV GEN book DAT woman NOM man
'The man gave the woman a book.' *V-DO-IO-S*

b. Nagbigay ng libro ang lalaki sa babae *V-DO-S-IO*

c. Nagbigay ang lalaki ng libro sa babae *V-S-DO-IO*

d. Nagbigay ang lalaki sa babae ng libro *V-S-IO-DO*

e. Nagbigay sa babae ang lalaki ng libro *V-IO-S-DO*

f. Nagbigay sa babae ng libro ang lalaki *V-IO-DO-S*

2.3 Morphological case and pronouns

There are three particles that precede nominals in Tagalog: *ang* [aŋ], *ng* [naŋ], and *sa* [sa]. Any given nominal can only be preceded by at most one of these particles. These are

often analyzed as morphological case.[2] In general, clauses have one argument that is singled out as the *ang*-marked argument, which I label as nominative. The *ang*-marked argument is obligatorily interpreted as specific (Rackowski & Richards, 2005). Pre-theoretically, it may bear any of a number of thematic relations to the verb, given the appropriate morphology on the verb (Schachter & Otanes, 1983). See the section below (Section 2.5) on voice morphology. Other post-verbal elements are either marked with *ng* or *sa*: *sa* is generally used for goals, recipients, and locations, while *ng* is used for possessors, instruments, and indefinite objects (Kroeger, 1993).

Ang, ng and *sa* exhibit allomorphy depending on whether the noun they are associated with is a common noun or a proper noun. This allomorphy is provided in Table 2.1.

Table 2.1: Case marker allomorphy

	NOM	GEN	DAT
Common noun	*ang*	*ng*	*sa*
Proper noun	*si*	*ni*	*kay*

Pronouns also exhibit this tripartite distinction, as provided in Table 2.2. Nominative pronouns have two forms: a second-position clitic and a free form used pre-verbally or in focus positions (e.g., as pivots of clefts). The clitic and the free form are the same, except for in the second-person singular where the clitic is *ka* and the free form is realized as *ikaw*. Genitive pronouns only have second-position clitic forms. Dative pronouns are free forms and they almost always appear with *sa* or *kay*. See Kaufman (2010) for a more detailed treatment of the pronominal paradigms in Tagalog and Kroeger (1998) for an overview of the second-position in Tagalog and the elements that can occupy this position.

[2]There is no consensus in the literature with regards to how *ang*, *ng*, and *sa* are glossed: Kroeger (1993) refers to them as NOMINATIVE, GENITIVE, and DATIVE, respectively; Rackowski and Richards (2005) as ANG, CASE, and DATIVE; Sabbagh (2011) as SUBJECT, NON-SUBJECT, and OBLIQUE; and Aldridge (2012) as ABSOLUTIVE, ERGATIVE/GENITIVE, and DATIVE. For ease of exposition, I adopt the glossing convention used by Kroeger (1993) throughout and will modify examples from other sources for consistency.

Table 2.2: Pronominal allomorphy

	NOM	GEN	DAT
1SG	*ako*	*ko*	*akin*
2SG	*ikaw/ka*	*mo*	*iyo*
3SG	*siya*	*niya*	*kaniya*
1PL.EXCL	*kami*	*namin*	*amin*
1PL.INCL	*tayo*	*natin*	*atin*
2PL	*kayo*	*ninyo*	*inyo*
3PL	*sila*	*nila*	*kanila*

As we will see later in Section 2.7.3, Aldridge (2002) beautifully uses the second-position requirement of certain pronouns to argue for the different structures of argument *wh*-questions and adjunct *wh*-questions in the language.

2.4 An excursus: Subjecthood in Tagalog

Subjecthood in Tagalog—or perhaps more generally, Philippine-type languages[3]—still remains controversial, even among Austronesiansts, because there is a bifurcation of the syntactic properties that are cross-linguistically controlled by the argument referred to as the subject in more familiar languages. Some of these properties are controlled by the agent, and some, by the *ang*-marked argument. Kroeger (1993) provides a detailed description of this bifurcation in Tagalog, and I summarize his description in Table 2.3.

This bifurcation led to different views in terms of how grammatical relations map onto thematic relations. Let's consider the following examples below:

(3) Adapted from W. Foley and Van Valin, Jr. (1984, p. 135)

[3]The term *Philippine-type* is based on typology. There are a number of languages that share this elaborate verbal morphology that is characteristic of central Philippine languages. It is not based on geography, as this classification extends to the Austronesian languages spoken in Taiwan, Borneo, Sulawesi, and Madagascar, and excludes some of the languages spoken in the Philippines (Kroeger, 2010),

a. Bumili ang **lalaki** ng isda sa tindahan
bought.AV NOM man GEN fish DAT store
'The man bought fish at the store.'

b. Binili ng lalaki ang **isda** sa tindahan
bought.PV GEN man NOM fish DAT store
'The man bought the fish at the store.'

Table 2.3: The bifurcation of subject-properties in Tagalog

	Ang-marked argument	Agent
1	It is the obligatory element of a clause	The agent, irrespective of case, is always a possible antecedent for a reflexive in the same clause
2	When quantifiers float, they are interpreted as referring to the nominative argument	It is typically the target of control constructions, modulo certain predicates and verbs in the non-volitive mood
3	It controls the optional number agreement in adjectives and in verbs	It is the addressee of imperatives

Early grammars referred to the *ang*-marked argument as the subject (Bloomfield, 1917, *inter alia*). Under this view, the example in (3-a) is considered a basic transitive sentence: the *ang*-marked *lalaki* "man" is the subject of the sentence, while the *ng*-marked *isda* "fish" is the direct object. Meanwhile, the example in (3-b) is an intransitive sentence, namely a passive construction: the *ang*-marked *isda* "fish" is now the subject of the sentence. This choice of labels, however, implies that (3-b) is more marked than (3-a), just as passives are in English or other more familiar languages. However, clauses of type (3-b) do not seem to be more marked. In fact, corpus studies (Cooreman, Fox, & Givón, 1984; Garcia et al., 2019; Nagaya, 2019) have shown that (3-b) is more common than (3-a).

McKaughan (1958) departed from the Bloomfieldian tradition of referring to the *ang*-marked argument as the subject in his description of Maranao, another Philippine-type language spoken in the Philippines. He reserved the term subject for the agent of the

clause, and instead, used "topic" for the *ang*-marked argument. It is highly unlikely that McKaughan had a full conception of what is now referred to as symmetric or Austronesian alignment (W. Foley, 1998, *inter alia*) in the typology literature, but his observations are important in that they acknowledged the difference mentioned above between these "passives" and typical passives in Indo-European languages. Some scholars, influenced by the observations of McKaughan (1958), have analyzed the examples in (3-a) and (3-b) as both being transitive sentences, suggesting that neither is more marked than the other.

Other researchers have analyzed Philippine-type languages as having an ergative system (Aldridge, 2004b, 2012, *inter alia*). Under this view, the example in (3-b) is a basic transitive sentence: the *ng*-marked argument *lalaki* "man" is the subject and the *ang*-marked argument *isda* "fish" is the object. Meanwhile, the example in (3-a) is an anti-passive: the *ang*-marked argument *lalaki* "man" is the subject. See V. Chen (2017) for arguments against an ergative analysis of Tagalog.

It is outside the scope of this dissertation to provide a more detailed treatment of the different views of subjecthood in Tagalog and ultimately, its morphological alignment. Throughout, I use the term subject (begrudgingly) to mean the agent of a semantically transitive predicate, irrespective of its case marking. More concretely, I treat the agent *lalaki* "man" as the subject, and the patient *isda* "fish" as the direct object in (3-a) and (3-b). I fully acknowledge that this a gross simplification of the facts and is primarily for expository ease and readability.

2.5 Voice morphology

Tagalog verbs typically carry what is often called VOICE MORPHOLOGY, which always cross-references the nominative argument of the clause, which is marked by *ang* [aŋ]. The argument that voice morphology cross-references will be in bold and the subject (=agent) will be underlined throughout.

Consider the sentences in (4). They have the same event structure—a man, a fish, and a store are involved in a buying event, and the thematic relations assigned by the verb to its arguments remain constant throughout. Notice that the verb form changes and the case markers of the arguments vary depending on the verb form. In (4-a), the verb exhibits AGENT VOICE (AV). The subject *lalaki* 'man' is marked nominative; the object *isda* 'fish', genitive, with *ng* [naŋ]; and the location *tindahan* 'store', dative, with *sa*. In (4-b), the verb exhibits PATIENT VOICE (PV). The object is now marked nominative; the subject, now genitive; and the location, still dative. In (4-c), the verb exhibits LOCATIVE VOICE (LV). The location 'store' is now marked nominative, and the other arguments, genitive.

(4) Adapted from W. Foley and Van Valin, Jr. (1984, p. 135)

a. Bumili ang **lalaki** ng isda sa tindahan
bought.AV NOM man GEN fish DAT store
'The man bought fish at the store.' *Agent voice (AV)*

b. Binili ng lalaki ang **isda** sa tindahan
bought-PV GEN man NOM fish DAT store
'The man bought the fish at the store.' *Patient voice (PV)*

c. Binilhan ng lalaki ng isda ang **tindahan**
bought.LV GEN man GEN fish NOM store
'The man bought fish at the store.' *Locative voice (LV)*

Following Rackowski (2002) and Aldridge (2004b), I assume that LV and the other voices are instances of high applicatives à la Pylkänen (2001). See V. Chen (2017) for arguments against an applicative analysis of these voices, however.

2.6 The interaction of voice and word order

The order of the elements after the verb is relatively free (Schachter & Otanes, 1983). However, some word orders are more natural than others. Kroeger (1993) maintains that

there are three competing pressures that can affect post-verbal word order:

(5) Adapted from Kroeger (1993, p. 111)

a. SUBJECT-FIRST: the subject (=agent) tends to precede all other arguments

b. NOMINATIVE-LAST: the *ang*-marked argument tends to follow other arguments

c. HEAVY NP SHIFT: "heavier" arguments tend to follow "lighter" arguments

First, consider the examples in (6). When the verb has PV, VSO, as in (6-a), is considered to be more natural than VOS, as in (6-b). VSO is preferred because it satisfies *both* of subject-first and nominative-last pressures, holding heaviness constant.

(6) The most natural word order for verbs with PV is VSO

a. Iniinom ng babae ang **alak**
drink.PV GEN woman NOM wine
'The woman is drinking the wine.' *PV with V-O-S order*

b. ?Iniinom ang **alak** ng babae
drink.PV NOM wine GEN woman
'The woman is drinking the wine.' *PV with V-S-O order*

Now consider the examples in (7). When the verb has AV, both VSO and VOS order are considered to be natural. No word order is optimal since both word orders satisfy one pressure at the expense of another. More specifically, with VSO, subject-first is satisfied at the expense of nominative-last. With VOS, it is the reverse: nominative-last is satisfied at the expense of subject-first.

(7) The most natural word order for verbs with AV is either VSO or VOS

a. Umiinom ang **babae** ng alak
drink.AV NOM woman GEN wine
'The woman is drinking wine.' *AV with V-S-O order*

b. Umiinom ng alak ang **babae**
drink.AV GEN wine NOM woman
'The woman is drinking wine.' *AV with V-O-S order*

There is experimental evidence suggesting that comprehenders are sensitive to these pressures. Bondoc and Schafer (2019) conducted a production study, where they gave participants the beginning of a sentence containing a verb with voice morphology, and asked them to complete it using the "ingredients" provided to them in the visual scene. For example, participants were supplied with *niluto kani-kanina lang* ('...just cooked earlier,' literal: cooked.PV earlier just) and were shown a picture of a little girl, spaghetti, and a pan. They found that when the verb had AV, about half of the time, participants filled the first position after the verb with the subject (i.e., the little girl) and the other half, with the object (i.e., spaghetti). Bondoc and Schafer interpreted this as being consistent with VSO or VOS word order being natural word orders when the verb has AV. However, when the verb had PV, participants overwhelmingly filled the first position with the agent. Bondoc and Schafer interpreted this as being consistent with VSO being the most natural word order when the verb has PV.

This discussion about the interaction between voice and word order will become relevant later in Chapter 5. As we will see, comprehenders prefer relative clauses with subject-gaps (SRCs; e.g., *The reporter that attacked the senator*...) over relative clauses with object-gaps (ORCs; e.g., *The reporter that the senator attacked*...). One account for this well-known asymmetry appeals to the similarity between the word order of a language's main clause and the order of the elements involving a relative clause in that language (Bever, 1970; Diessel & Tomasello, 2005). In English, for example, the basic word order is SVO. The order in a transitive SRC is also SVO, while the order in an ORC is OSV. Thus, under this account, the asymmetry obtains because the SRC order is similar to the basic word order. In Chapter 5, I use Tagalog as a case study to evaluate the empirical coverage of the

various accounts proposed in the literature to account for the asymmetry.

2.7 Three syntactic constructions in Tagalog

In this section, I provide an overview of three syntactic constructions in Tagalog that are featured in this dissertation: *ay*-inverted sentences, relative clauses, and *wh*-questions. All three constructions are featured in the judgment studies presented in Chapter 3 and the stops-making-sense experiments in Chapter 4. Relative clauses are also featured in the picture-selection experiments in Chapter 5.

2.7.1 *Ay*-inversion

Recall that in Tagalog, verbs typically come first. Schachter and Otanes (1983) describe different types of inversion processes in the language that change the word order by fronting one of the post-verbal elements. Here I only discuss one of them, *ay*-inversion.

This process changes the word order by fronting a constituent and adding the morpheme *ay* in between the verb and this constituent, which may be the nominative argument, as in (8-a); an adverbial, as in (8-b) and (8-c); or a gerund, as in (8-d).

(8) Adapted from Kroeger (1993, pp. 65, 67–8)

a. **Siya** ay nagbabasa ng diyaryo
3SG.NOM AY read.AV GEN newspaper
'S/he is reading a newspaper.' *Nominative argument*

b. Madalas ay pumupunta **siya** dito
often AY go.AV 3SG.NOM here
'He comes here often.' *Manner adverb*

c. Bukas ay **kami**-y pupunta
tomorrow AY 1PL.NOM.EXCL-AY go.PV
'We'll go tomorrow.' *Temporal adverb*

d. Pagdating ko sa Pilipinas ay nagpunta **ako** sa Baguio
arriving 1SG.GEN DAT Philippines AY went.AV 1SG.NOM DAT PN
'Upon arriving in the Philippines, I went to Baguio.' *Gerund*

The morpheme *ay* has two surface forms depending on its phonological environment: (i) the enclitic form, which is written as *'y* and pronounced as [j], is used when the preceding word ends in a vowel or a glottal stop; and (ii) the free form, which is written as *ay* and pronounced as [aj], may be used in all contexts.

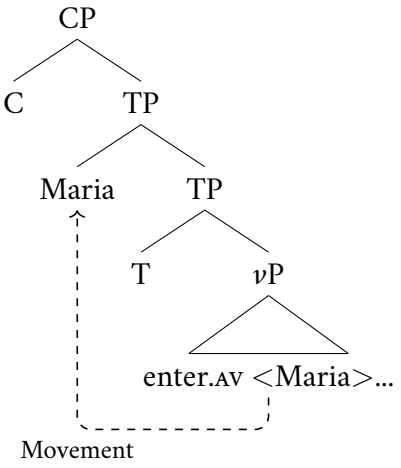
In terms of usage, Schachter and Otanes (1983, p. 486) describe this operation as completely optional, with no obvious effect on the meaning of the sentence. Stylistically, this is characteristic of more formal style, and is more common in writing than in spoken Tagalog. On the basis of pragmatics, Kroeger (1993) argues that pre-verbal position in these constructions is a grammaticized topic position because it is incompatible with focus.

In terms of its structure, Sabbagh (2014) argues that the fronted nominal, like *si Maria* in (9-a), is in the specifier position of some functional projection FP that is in between T and C, as schematized in (9-b). The main evidence he provides is that *ay*-inverted sentences can be embedded under a verb like *tanong* 'ask,' as shown in (9-c). What is crucial about the structure of *ay*-inverted sentences is that the relationship between the fronted noun and the position where it is interpreted is a direct dependency. In other words, it is a direct A-bar chain produced by movement.

(9) Adapted from Sabbagh (2014)

a. Si **Maria** ay lumulusong sa bayan upang magtinda
NOM PN AY enter.AV DAT town in.order sell.AV
'Maria goes to town in order to sell things.'

b.



c. Itanong sa sarili mo [$_{CP}$ kung ang bibilhin mo ay isa-ng need
ask.PV DAT self 2SG.GEN COMP NOM buy.PV 2SG.GEN AY one-LNK need
o isa lamang want]
or one only want
'Ask yourself whether what you're buying is a need or a want.'

2.7.2 Relative clauses

Tagalog is a typologically rare language (Dryer, 2013), in that it allows the head noun to surface in a variety of positions with respect to a relative clause modifier (Aldridge, 2017; Law, 2016). First, the head noun may surface before the RC, as in (10-a). Second, it may surface after, as in (10-b). Third, it may surface in a position that seems internal to the RC, as in (10-c). Finally, it may be null, as in (10-d).

(10) Adapted from Aldridge (2017)

a. ang **libro**-ng binili ng babae
NOM book-LNK bought.PV GEN woman
'the book that the woman bought' *Head-initial*

b. ang binili ng babae-ng **libro**
NOM bought.PV GEN woman-LNK book
'the book that the woman bought' *Head-final*

c. ang binili-ng **libro** ng babae
NOM bought.PV-LNK GEN woman
'the book that the woman bought' *Head-internal*

d. ang binili ng babae
NOM bought.PV GEN woman
'what the woman bought' *Headless*

Head-initial RCs are featured in the judgment studies presented in Chapter 3, the stops-making-sense experiments in Chapter 4, and the picture-selection task in Chapter 5. Head-final RCs are also featured in Chapter 5. As we will see, headless RCs are also featured in the judgment studies presented in Chapter 3 and the stops-making-sense experiments in Chapter 4 via argument *wh*-questions. As we will see below, argument *wh*-questions are analyzed as a (pseudo-)cleft where the *wh*-word is a predicate nominal and the remaining material is inside a headless RC (Aldridge, 2002; Kroeger, 1993).

Let's discuss some properties of RCs that may become relevant later on. First, although the language has these four logical options available for the word order of a head noun and a RC modifier, they are not all equal in terms of markedness and/or naturalness. My work with speakers indicate that head-initial RCs are the unmarked and most natural sounding option. Head-final and head-medial RCs—especially those with referential expressions as co-arguments—are more marked and not as natural sounding. In general, speakers prefer RCs that are "smaller" in size, which could mean that the relative clause contained an intransitive verb, an existential, or a transitive verb with pronominalized or null arguments. When the RCs are smaller in size, as in (11), speakers judge head-initial and head-final RCs as equally natural sounding.

(11) Constructed examples

a. ang **kendi**-ng binili niya
NOM book-LNK bought.PV 3SG.GEN
'the candy that s/he bought' *Head-initial*

b. ang binili niya-ng **kendi**
NOM bought.PV 3SG.GEN-LNK candy
'the candy that s/he bought' *Head-final*

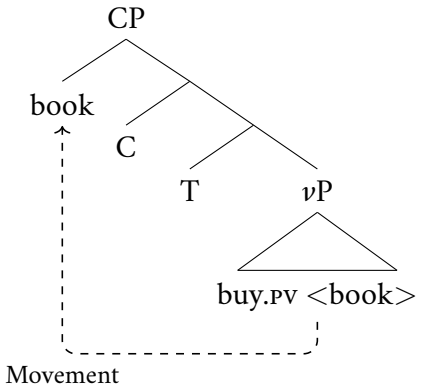
Second, they are also not equal in terms of frequency. Nagaya (2019) conducted a corpus analysis of naturally occurring conversations and found that headless RCs, as in (10-d), were the most frequent, accounting for 80% of instances of RCs. Among the headed RCs, he found that head-initial RCs, as in (10-a), were the most frequent, accounting for approximately 15% of the data, and head-final RCs, as in (10-b), 5% of the data. There was no mention of head-internal RCs, as in (10-c). It is unclear to me if this meant that there were no instances of head-internal RCs or if he did not code/include them in his analysis.

Third, another insight from Nagaya is that collapsing across RC-type, PV is more frequent than AV. PV occurred within 75% of the RCs in his corpus. This echoes the more general finding that PV is more frequent than AV in the language (Cooreman et al., 1984; Garcia et al., 2019). Lastly, in terms of structure, whether these RCs should share a single derivation remains an open question. See Law (2016) for an analysis that unifies the four types of RCs. See Aldridge (2004b, 2017) for alternative views. Despite the differences in how Law and Aldridge derive head-final and head-internal RCs, they agree that head-initial RCs are derived via head-raising, as schematized in (12). Thus, the relationship between the head noun and the position where it is interpreted is a direct dependency, a direct A-bar chain produced by movement, at least for head-initial relative clauses.

(12) Adapted from Aldridge (2017)

a. ang **libro**-ng binili ng babae
NOM book-LNK bought.PV GEN woman
the book that the woman bought'

b.



2.7.3 *Wh*-questions

Consider the two types of *wh*-questions in (13). The example in (13-a) is an argument *wh*-question, while the example in (13-b) is an adjunct *wh*-question. Superficially, both of them have the *wh*-word in clause-initial position, much like in their English counterparts. However, there are major differences to notice. First, the nominative case marker is present after the *wh*-word in (13-a) but absent in (13-b). Second, the placement of the second-position clitic *niya* varies. It is in its expected second-position in (13-b) but it is in a curious position in (13-a).

(13) Constructed example

a. **Ano** ang ininom niya kagabi
what NOM drank.PV 3SG.GEN last night
'What did s/he drink last night?' *Argument wh-question*

b. Bakit/Saan/Kailan niya ininom ang **alak** kagabi
why/where/when 3SG.GEN drank.PV GEN wine last night
'Why/where/when did s/he drink wine last night?' *Adjunct wh-question*

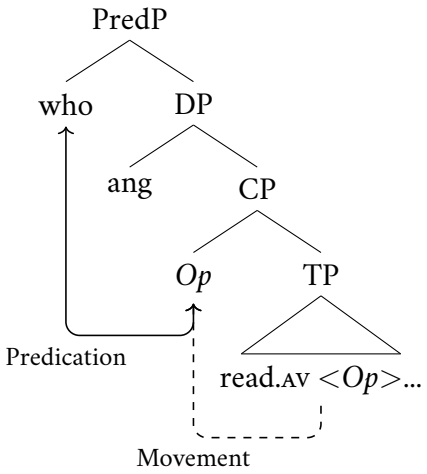
While it might be tempting to derive (13) the same way, there are multiple paths to deriving *wh*-questions cross-linguistically. See Potsdam (2009) for an overview of the different paths to deriving *wh*-initiality in Austronesian languages. In Tagalog, *wh*-questions are derived in two ways, depending on whether the *wh*-phrase is an argument or an adjunct.

When the *wh*-phrase is an argument, the language uses a (pseudo-)clefting strategy, where the interrogative phrase serves as the predicate nominal and the remaining material is a headless relative (Aldridge, 2002; Kroeger, 1993). When the *wh*-phrase is an adjunct, the language use the same mechanism used in English, *wh*-movement.

(14) Adapted from Aldridge (2002)

a. **Sino** ang bumabasa ng libro?
who NOM read.AV GEN book
'Who is reading a book?'

b.



Now let us consider how Aldridge (2002) motivated these different derivations. She noticed that the predicate nominal in (15-a) and the *wh*-word (15-b) are in complementary distribution. The example in (15-a) is standardly analyzed as follows: *kabastusan* is a predicate nominal and the constituent *ang ginagawa mo* is a headless RC. If we assume that *ano* is also a predicate nominal via complementarity, then the presence of the nominative marker *ang* after the *wh*-word is expected.

(15) Constructed examples

a. **Kabastusan** ang ginagawa mo
Vulgarity NOM do.PV 2SG.GEN
'The thing you're doing is vulgarity.' *Predicate nominal*

b. **Ano** ang ginagawa mo
what NOM do.PV 2SG.GEN
'What are you doing?'

Literal: The thing you're doing is what? *Argument wh-question*

This analysis of argument *wh*-questions also allows us to explain the curious behavior of the second-position clitic *niya* in (13), repeated here as (16). When the clausal boundaries (which are demarcated using "[ ]") are considered, the recalcitrant behavior of *niya* in (16-a) is now only apparent: *niya* is in second-position in its clause (i.e., inside the headless RC).

(16) Constructed example

a. [$_{CP}$ **Ano** ang [$_{CP}$ ininom niya kagabi]]
what NOM drank.PV 3SG.GEN last night
'What did s/he drink last night?' *Argument wh-question*

b. [$_{CP}$ Bakit/Saan/Kailan niya ininom ang **alak** kagabi]
why/where/when 3SG.GEN drank.PV GEN wine last night
'Why/where/when did s/he drink wine last night?' *Adjunct wh-question*

These two ways of deriving *wh*-initiality mean that in argument *wh*-questions, the relation between the *wh*-word and position where it is interpreted is an indirect one. Although the null operator inside the headless RC does undergo *wh*-movement, the way in which it becomes associated with *wh*-word is mediated via predication.

Before proceeding, as we will see in Chapter 4, *d*-linked *wh*-questions, which are *wh*-questions that have the shape *which N* in English, also figure in this dissertation. I learned after conducting Experiment 5 that there are (at least) two types of *d*-linkers in Tagalog. Some speakers use *alin* 'which' as a catch-all *d*-linked interrogative, as in (17-a), irrespective of the animacy of the noun being *d*-linked. During debriefing of Experiments 4–5 (and subsequent work with consultants), I learned that other speakers have a more restricted environment for *alin*. These speakers can only use *alin* with inanimate nouns. In

order for an animate noun to be *d*-linked, *sino* 'who' needs to be used. For example, these speakers would use *sinong bata* 'which child' instead of *aling bata*, as in (17-b). One of my language consultants commented that this might be prescriptivism in action.

(17) Two types of *d*-linkers in Tagalog

a. Alin-g bata ang pasaway?
which-LNK child NOM pesky
'Which child is pesky?' *D-linker 1: Alin + animate noun is okay*

b. Sino-ng/#Alin-g bata ang pasaway?
who-LNK/which-LNK child NOM pesky
'Which child is pesky?' *D-linker 2: Alin + animate noun is not okay*

2.8 The interaction of voice and extraction

As in many other Austronesian languages, Tagalog voice morphology interacts with movement dependencies, like the constructions surveyed above. This interaction has been called many names in the generative syntax literature, from the "subject-only restriction" (Keenan & Comrie, 1977) to the "absolutive restriction on Ā-extraction" (Aldridge, 2008). To be more descriptive, I refer to this interaction throughout as the "extraction restriction." The received generalization that emerges upon surveying the Tagalog syntax literature is provided in (18). I refer to the received view as VOICE MATCH EXTRACTIONS—to contrast with what I will be referring to as VOICE MISMATCH EXTRACTIONS and preview the experimental manipulations in the judgment studies presented in Chapter 3.

(18) The *voice match* extraction restriction in Tagalog

The *ang*-marked argument (*i.e.,* the argument cross-referenced by voice) is the only argument eligible to participate in movement dependencies (Schachter, 1977; Ceña, 1979; Aldridge, 2002; Rackowski & Richards, 2005; Kaufman, 2009; Law,

2016; Aldridge, 2017, 2018).

What (18) says is that an argument that participates in movement dependencies must match the argument cross-referenced by voice. In (19), I show how voice interacts with extraction. When the verb exhibits AV, the subject *babae* 'woman,' which is cross-referenced by AV, is extractable and can serve as the pivot of a cleft, as in (19-a). Meanwhile, the object *baro* 'dress,' which is not cross-referenced by AV, is not extractable and cannot serve as the pivot of a cleft, as in (19-b). In contrast, when the verb exhibits PV, the object 'dress,' which is cross-referenced by PV, is extractable and can serve as the pivot of a cleft, as in (19-c). The subject 'woman,' which is not cross-referenced by PV, is not extractable and cannot serve as the pivot of a cleft, as in (19-d).

(19) Adapted from Schachter (1977, p. 286)

a. Iyon ang **babae**-ng bumili ng baro
that NOM woman-LNK bought.AV GEN dress
'That's the woman who bought a dress.' *Voice match with AV*

b. *Iyon ang baro-ng bumili ang **babae**
that NOM dress-LNK bought.AV NOM woman
Intended: That's the dress that the woman bought. *Voice mismatch with AV*

c. Iyon ang **baro**-ng binili ng babae
that NOM dress-LNK bought.PV GEN woman
'That's the dress that a/the woman bought.' *Voice match with PV*

d. *Iyon ang babae-ng binili ang **baro**
that NOM woman-LNK bought.PV GEN dress
Intended: That's the woman who bought the dress. *Voice mismatch with PV*

For ease of exposition, I exemplified voice match restriction using AV and PV. The restriction provided in (18) extends to the other voices in the language as well (Rackowski & Richards, 2005). For example, when the verb exhibits LV, as in (4-c), only the location *tindahan* 'store,' which is cross-referenced by LV, is extractable and can serve as the pivot

of a cleft, as in (20-a). The other arguments, on the other hand, are not extractable and cannot serve as the pivot of a cleft, as in (20-b)–(20-c), respectively.

(20) Constructed example

a. Iyon ang **tindahan**-g binilhan ng lalaki ng isda
that NOM store-LNK bought.LV GEN man GEN fish
That's the store where the man bought fish. *Voice match with LV*

b. *Iyon ang lalaki-ng binilhan ang **tindahan** ng isda
that NOM man-LNK bought.LV NOM store GEN fish
Intended: That's the man who bought fish at the store. *Voice mismatch*

c. *Iyon ang isda-ng binilhan ng lalaki ang **tindahan**
that NOM fish-LNK bought.LV GEN man NOM store
Intended: That's the fish the man bought at the store. *Voice mismatch*

2.9 The theoretical landscape

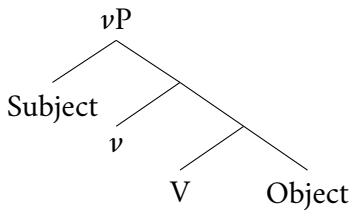
Here I survey four recent approaches to voice morphology and how they account for voice match extractions in Tagalog. Despite the differences in the finer details of these proposals, they can be (roughly) classified into two classes of analyses: (i) those involving multiple specifiers; and (ii) those that treat voice morphology as agreement with C.

2.9.1 Approach 1: Multiple specifiers

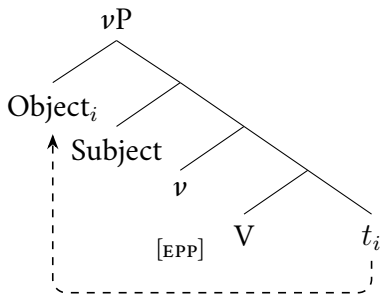
Analyses under the umbrella of "multiple specifiers" have these features in common. First, AV and non-AV voices are structurally different. In non-AV voices (e.g., PV, LV, etc), the highest internal argument vacates the VP and moves to the outermost specifier of *v*P, above the subject. Second, the highest DP in the specifier of *v*P is spelled out as *ang*. Third, voice match extraction, as will be discussed below, is accounted for by appealing to a phase-based theory of extraction, supplemented by some version of Attract Closest.

Within this approach, there are two analyses of voice morphology. Let us call one version the CASE-AGREEMENT analysis (Rackowski & Richards, 2005; Sabbagh, 2016, *inter alia*). Let us call the other one the TRANSITIVITY analysis (Aldridge, 2008, 2012, *et seq.*). Before discussing the specifics of each, let's consider their main similarity. Under both analyses, AV and PV have different derivations, as shown in (20-a) and (20-b). PV involves moving the object to the outermost specifier of *v*P to satisfy the [EPP]-feature on *v*.

(21) Partial derivation of AV



(22) Partial derivation of PV

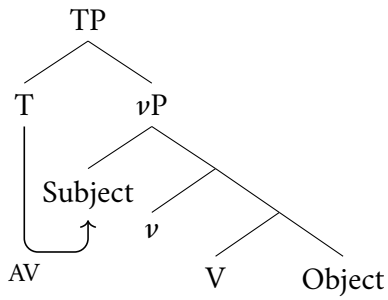


The motivation for this movement differs, however. Under a Case-agreement analysis, PV is akin to object shift in Germanic languages (Collins & Thráinsson, 1996; Vikner, 2005). The object vacates the VP and moves to the outermost specifier of *v*P in order to be interpreted as specific. Under a transitivity analysis, the movement of the object to the outermost specifier of *v*P is a consequence of merging a flavor of *v*—transitive *v*—that obligatorily carries an [EPP]-feature, which forces the object to move to its specifier.

Now let's consider the two analyses in greater detail. Under the Case-agreement analysis, voice morphology is the reflex of the Agree relation between T and the highest DP in the specifier of *v*P. This Agree relation does not involve the familiar type of ϕ-agreement,

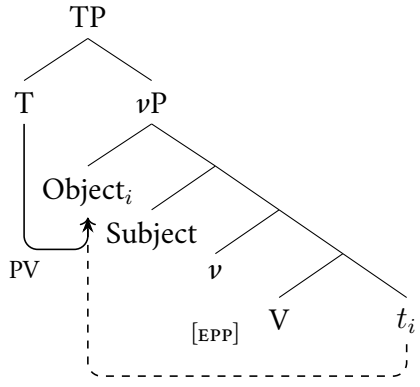
but instead, abstract Case-agreement. Let's first consider the derivation when verbs exhibit AV. When T is merged, it probes down and finds the subject DP occupying the outermost specifier of *v*P. T Agrees with the abstract Case value of the subject and the morphological reflex of this Agree relation is AV. This is schematized in (23).

(23) Partial derivation of AV



Now consider the derivation when the verb exhibits PV. As discussed above, the object vacates the VP and moves to the outermost specifier of *v*P to be interpreted as specific. When T is merged, it probes down and finds the object DP occupying the outermost specifier of *v*P. T Agrees with the abstract Case value of the object and the morphological reflex of this Agree relation is PV. This is schematized in (24).

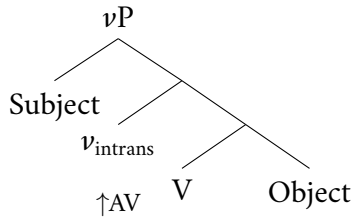
(24) Partial derivation of PV



Under a transitivity analysis, voice morphology is the spell-out of two different *v*'s. PV is the spell-out of transitive *v*, while AV is the spell-out of intransitive *v*. The main

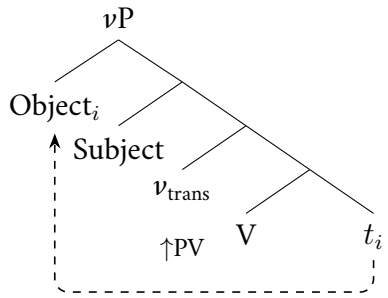
difference between the two *v*'s is their featural make-up: transitive *v* obligatorily has an [EPP]-feature, while the intransitive *v* does not have it. Crucial to this analysis is an ergative view of Tagalog, where sentences with PV are basic transitives and sentences with AV are intransitives—and more specifically, anti-passives. Let's first consider the derivation when verbs exhibit AV. The intransitive *v* is merged. This flavor of *v* does not carry an [EPP]-feature and thus, the object does not evacuate the VP. When T is merged, it assigns structural case on the subject DP and the object DP receives lexical case from V. This is schematized in (25).

(25) Partial derivation of AV



Now let's consider the derivation when the verb exhibits PV. When the transitive *v* is merged, its [EPP]-feature forces the object to vacate the VP and move to the outermost specifier of *v*P. When T is merged, it assigns structural case on the object. The subject DP receives inherent case from *v*. This is schematized in (26).

(26) Partial derivation of PV



Finaly, let's consider how these analyses account for the interaction between voice and

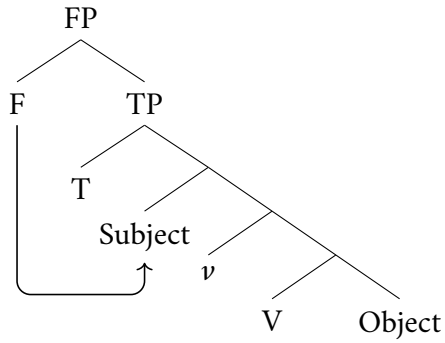
extraction. They both enforce the voice match extraction restriction by appealing to a phase-based theory of extraction (Chomsky, 2000, 2001), where an element XP can be extracted out of a phase if and only if it is located at the edge of the phase. They need to supplement this with some version of Attract Closest, where closest is defined in (27):

(27) From Rackowski and Richards (2005, p. 579)

"A goal α is the closest one to a given probe if there is no distinct goal β such that for some X (X a head or maximal projection), X c-commands α but does not c-command β."

With this formal machinery in place, the voice match extraction restriction is a natural consequence. As schematized in (28), when the verb exhibits AV, only the subject DP is at the phase-edge. Thus, if a probe *F* were to search for a goal, under a phase-based theory of extraction, only the subject is visible to that probe because it is the most structurally prominent DP, by virtue of being the only one in the specifier of *v*P. On the other hand, the object is opaque to this probe. This captures the desired patterns of (un)grammaticality.

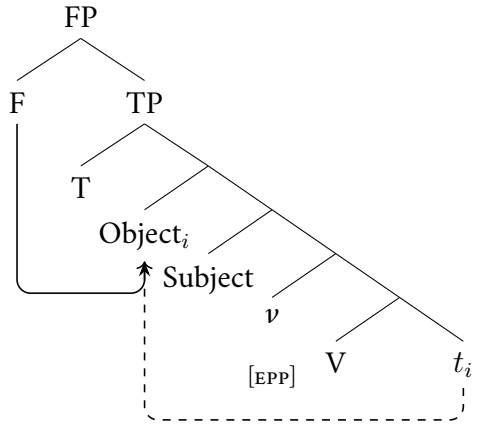
(28) Partial derivation of AV



As schematized in (29), when the verb exhibits PV, the object evacuates the VP and moves to the outermost specifier of *v*P. Thus, if a probe F were to search for a goal, under a phase-based theory of extraction *plus* some version of Attract Closest (27), the object is the

more prominent DP and is thus the closest goal visible to that probe. On the other hand, the subject is no longer the most prominent DP at the edge of the *v*P-phase, rendering it opaque to this probe. This captures the desired patterns of (un)grammaticality.

(29) Partial derivation of PV



2.9.2 Approach 2: Agreement with C

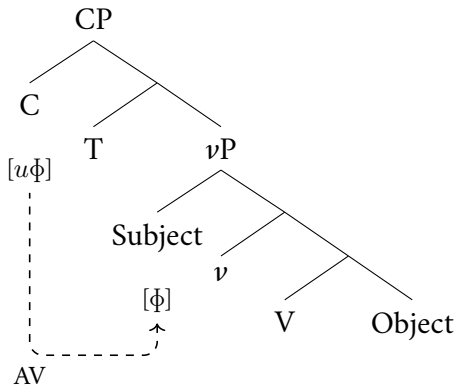
Analyses under the umbrella of "agreement with C" have the following features in common. First, voice morphology is the spell-out when C enters into an Agree relation with whatever DP surfaces with *ang*. Second, voice match extraction is captured by singling out C as the only driver of movement in the language. Within this approach, there are two analyses of voice. Let us call one version ϕ-AGREEMENT analysis (Aldridge, 2018). Let us call the other one the BUNDLING analysis (V. Chen, 2017)

Under a ϕ-agreement analysis, voice morphology is the spell-out of ϕ-agreement when C, which has an unvalued feature, say [*u*ϕ], probes down, sees a DP that has a feature [ϕ] that could value [*u*ϕ], and Agrees with that DP. Voice morphology is the spell-out of this Agree relation. Aldridge (2018), In short, AV is ϕ-agreement with the subject; PV is ϕ-agreement with the object.

Let's first consider the derivation when the verb exhibits AV. When C probes down its search space to value its unvalued feature, [*u*ϕ], it sees the subject with the feature [ϕ]

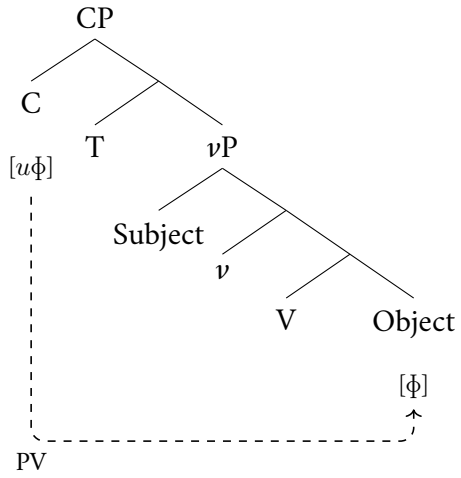
that could value [$u\phi$]. C Agrees with this DP and the result is AV. In other words, AV is ϕ-agreement with the subject. This is schematized in (30).

(30) Partial derivation of AV



Now consider the derivation when the verb exhibits PV. When C probes down its search space to value its unvalued feature, [$u\phi$], it skips over the subject because *v* had already assigned inherent case via spec-head configuration. It continues searching and sees the object with the feature [ϕ] that could value [$u\phi$]. C Agrees with this DP and the result is PV. In other words, PV is ϕ-agreement with the object. This is schematized in (31).

(31) Partial derivation of PV



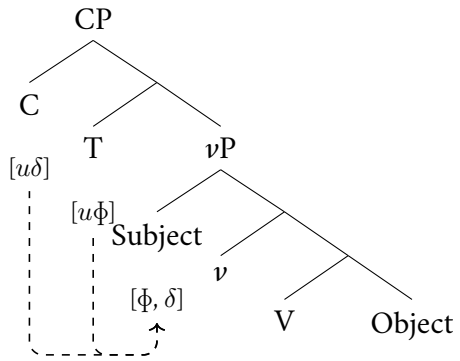
Under a bundling analysis, voice morphology is the spell-out of bundles of different

Agree relations targeting the same goal. In other words, voice morphology is the spell-out when C, which has a [$u\delta$], and another head, which has a [$u\phi$], Agree with the same DP and use its features to value [$u\delta$] and [$u\phi$], respectively.

Before proceeding, let's make explicit the two assumptions that this analysis makes. First, this analysis assumes that "Agree relations may bundle in their morphological output when more than one Agree relations targets the same goal" (V. Chen, 2017, p. 163). Second, this analysis assumes that object agreement is present in all human languages, even when there is no overt spell-out (V. Chen, 2017, p. 161). The formal implementation of object agreement is as follows: first, *v* (or in Chen's more articulated verbal spine, the Voice head), which has a [$u\phi$]-feature, probes down its search space; second, it finds the object; and third, it uses the [ϕ]-feature of the object to value its features.

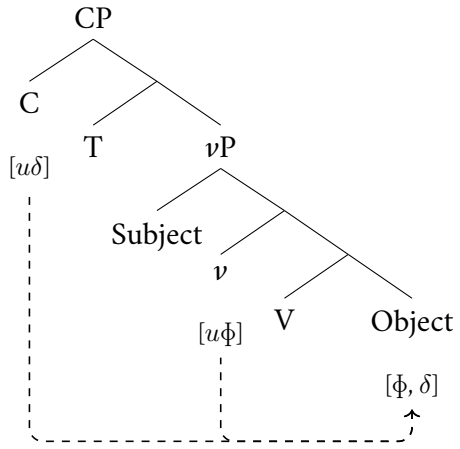
Let's consider the derivation when the verb exhibits AV. AV is the spell-out when C, which has a [$u\delta$], and T, which has [$u\phi$], both target the same goal, the subject, to value their respective features. This is schematized in (32).

(32) Partial derivation of AV



Now consider the derivation when the verb exhibits PV. PV is the spell-out when C, which has a [$u\delta$], and *v*, which has [$u\phi$], both target the same goal, the object, to value their respective features. This is schematized in (33).

(33) Partial derivation of PV



Both types of analyses enforce the voice match extraction by singling out C, which has an [EPP], as the only driver of movement in the language, as in seen in (34). If the DP also checks the [EPP]-feature during this Agree-relation, then the DP moves to the specifier of CP. Framed differently, movement is only possible for the DP with which C has entered into an Agree relation. This effectively enforces the voice match restriction.

(34) C is the only driver of A-bar movement

a. "[There is] only one feature driving movement to [Spec, CP], which is [uϕ]" (Aldridge, 2018)

b. "Types of A'-feature, such as [Top], [Rel], and [Foc], may be driven by a single "catch-all" A'- probe, as opposed to those in other languages, which employ separate probes for [Top], [Rel], and [Foc]." (V. Chen, 2017, p. 195)

With this formal machinery in place, the voice match extraction restriction falls out naturally. When the verb exhibits AV, C enters into an Agree relation with the subject and thus, the DP can be moved to the specifier of CP—if it also checks off C's [EPP]-feature. However, because C did not enter into an Agree relation with the object, that DP cannot be moved to the specifier of CP. Similarly, when the verb exhibits PV, C enters into an Agree relation with the object DP and thus, the DP can be moved to the specifier of

CP—if it also checks off C's [EPP]-feature. However, because C did not enter into an Agree relation with the subject, that DP cannot be moved to the specifier of CP.

An inquiring mind might want to know why some version of a locality principle—Attract Closest, for example—is not respected when C probes. More concretely, why does C "skip over" the subject and find the object instead when it probes? Aldridge (2018) argues that "it is no longer necessary to force or prevent movement to [Spec, *v*P] in order to ensure locality among DPs in this phase edge, since it is valuing of the ϕ-probe that determines which DP moves to [Spec, CP]." Framed differently, C "skips over" the subject when it probes because that DP is now licensed via inherent case from *v* and is now an inactive goal.

2.10 Conclusion

The main goal of this chapter is to provide an overview of the morphosyntactic facts of Tagalog that figure in the experiments that are presented in subsequent chapters. To reiterate, these are the main features that I want us to pay close attention to:

- Tagalog verbs carry voice morphology, which always cross-references the nominative argument of the clause (i.e., the *ang*-marked argument).

- Voice morphology interacts with word order. When the verb exhibits AV, both VSO and VOS are considered to be natural. When the verb exhibits PV, VSO is considered to be the most natural.

- In *ay*-inverted sentences, the relation between the fronted nominal and the position where it is interpreted is a direct dependency. In head-initial RCs, the relation between the head noun and the position where it is interpreted is also a direct dependency. Meanwhile, the relation between the *wh*-word and the position where

it is interpreted is not a direct dependency, but is rather mediated via predication. See the bullet point below.

- Argument *wh*-questions involve a bi-clausal structure, where the *wh*-word is a predicate nominal and the remainder of the sentence is inside a headless RC.
- Tagalog allows for both head-initial and head-final RCs.
- PV is more frequent than AV in the language more generally, and also in RCs, more specifically.
- The argument that is cross-referenced by voice is the only argument eligible to participate in movement dependencies.
- While there are different approaches to voice and the voice match extraction restriction, all of the analyses are movement-based.

Chapter 3

The extraction restriction, revisited

Tagalog voice morphology and its interactions with the other parts of the language's grammatical system are well-studied in the Austronesian literature. As discussed in Chapter 2, at first blush, Tagalog seems to exhibit all of the hallmark properties of a language with the Austronesian voice system (Erlewine, Levin, & van Urk, 2017; V. Chen & McDonnell, 2019):

- Verbal morphology that co-varies with a privileged argument;
- The privileged argument receives a particular morphological form irrespective of its original thematic role;
- Eligibility to participate in movement dependencies is often limited to this privileged argument.

This main goal of this chapter is to re-examine the nature of the interaction between voice morphology and A-bar movement in the language. I repeat the received empirical generalization from Chapter 2 concerning this interaction in (1).

(1) The *voice match* extraction restriction in Tagalog

Only the *ang*-marked argument (*i.e.,* the argument cross-referenced by voice) is

eligible for A-bar movement (Schachter, 1977; Ceña, 1979; Aldridge, 2002; Rackowski & Richards, 2005; Kaufman, 2009; Law, 2016; Aldridge, 2017, 2018).

What (1) says is that an argument that participates in movement dependencies must match the argument cross-referenced by voice. One of the main contributions of this chapter is to make available an alternative view of the extraction restriction: the eligibility to undergo A-bar movement is *not* limited to only the privileged argument. When subjects are not cross-referenced by voice, they still enjoy a certain degree of extractability—at least *for some speakers*. I provide naturally occurring data, as well as experimental data, consistent with this alternative view.

The chapter is organized as follows. Section 3.1 presents the alternative view. Section 3.2 provides naturally occurring data that are consistent with the alternative. Sections 3.3–3.5 present the results of a series of acceptability judgment studies that show the extractability of subjects even when they are not cross-referenced by voice. Section 3.6 develops quantitative models to better understand the underlying distributions of the participants' judgment process, using the framework developed by Dillon et al. (2017). Section 3.7 concludes by discussing the following: first, a reformulation of the competing generalizations as different grammars arising from competition during language acquisition; second, what the implications of the results are for syntactic theory; and finally, what the implications of the results are for sentence processing.

3.1 The alternative: Voice-mismatch restriction

Contra the generalization in (1), others have observed that the interaction may not be as strict and can allow extractions where the extracted nominal and the nominal cross-referenced by voice morphology do not match. More specifically, extractions of subjects are licit under NON-AV VOICES (NAV)—PV, LV, and the other voices in the language—

for some speakers (de Guzman, 1995; Ceña & Nolasco, 2011, 2012; Tanaka, 2016; Hsieh, 2019).[1] Throughout, I refer to the extraction restriction that is not as voice-sensitive as VOICE MISMATCH extraction restriction. Since the works by Ceña and Nolasco (2011, 2012) are written in Tagalog, this variation in the extraction restriction has remained largely undiscussed in mainstream generative linguistics.

The earliest attestation[2] of the voice mismatch restriction that I could find in the literature is the work by de Guzman (1995) that was couched in Relational Grammar.[3] She observed that subject may be topicalized when the verb exhibited PV, as in (2-a). The object, however, may not topicalize when the verb exhibited AV, as in (2-b).

(2) Adapted from de Guzman (1995, p. 57)

a. Ang nanay, lulutuin ang isda sa kusina
NOM mother will cook.PV NOM fish DAT kitchen
'The mother, she will cook the fish in the kitchen'

b. *Ang isda, magluluto ang nanay sa kusina
NOM fish will cook.AV NOM mother DAT kitchen
Intended: The fish, the mother will cook it in the kitchen

Ceña and Nolasco (2011, p. 181) provided more examples of the voice mismatch extraction. They observed the following contrasts in Tagalog. Note that I slightly modified their examples by adding *ang* to unambiguously mark the left edge of the RC. This becomes especially relevant in (6) when I talk about head-final RCs.

First consider examples where the verb exhibits AV, as in (3). They have noted that the

[1] Even though Hsieh (2019) treats agent extractions under PV as ungrammatical, he notes in footnote 10 that "some speakers do not judge agent extractions out of non-actor voices to be totally ungrammatical." In other words, they are not as well-formed as extractions where the extracted argument is cross-referenced by voice (i.e., agent-extractions under AV and patient-extractions under PV) but they are significantly better than patient-extractions under AV.

[2] There is even earlier work by Ceña (1979) that talks about possessor raising. In a sense, it is a type of voice mismatch extraction because what is being extracted is not cross-referenced by voice. However, I set sub-extraction aside since I am interested more in cases where the extracted argument is not a subpart of what is cross-referenced by voice but instead, a completely different argument.

[3] Many thanks to Nozomi Tanaka for bringing this to my attention.

subject *bantay* 'guard'—the argument cross-referenced by voice—can be relativized, as in (3-a). The object *basi* 'rice wine' cannot be relativized, as in (3-b). The source *tao* 'person' also cannot be relativized, as in (3-c). These patterns are consistent with the generalization in (1) thus far. Recall that the argument that the verb cross-references is be in bold, and the subject is underlined throughout.

(3) When the verb exhibits AV, only the subject can be relativized

a. ang **bantay** na bumili ng basi sa tao
NOM guard LNK bought.AV GEN rice wine DAT person
'the guard that bought rice wine from the person'

b. *ang basi na bumili ang **bantay** sa tao
ANG rice wine LNK bought.AV NOM guard DAT person
Intended: the rice wine that the guard bought from the person

c. *ang tao na bumili ng basi ang **bantay**
NOM person LNK bought.AV GEN rice wine NOM guard
Intended: the person that the guard bought rice wine from

The extraction facts reported by Ceña and Nolasco (2011, 2012) diverge from the generalization in (1) when the verb exhibits NAV. In (4), the verb exhibits PV. The object *tao* 'person'—the argument cross-referenced by voice—can be relativized, as in (4-a). The location *palengke* 'market' cannot be relativized—at least, not using the linker strategy, as in (4-b). It can, however, be relativized using a different relativization strategy. These patterns are consistent with (1) thus far. However, (4-c) is inconsistent with it: the subject *bantay* 'guard' can also be relativized.

(4) When the verb exhibits PV, the patient or the subject can be relativized

a. ang **tao** na niloko ng bantay sa palengke
NOM person LNK deceive.PV GEN guard DAT market
'thee person that the guard deceived at the market'

b. *ang palengke na niloko ng bantay ang **tao**
NOM market LNK deceive.PV GEN guard NOM person
Intended: the market where the guard deceived the person

c. ang bantay na niloko ang **tao** sa palengke
NOM guard LNK deceive.PV NOM person DAT market
'the guard that deceived the person at the market'

In (5), when the verb exhibits applicative morphology, the recipient *tao* 'person'—the argument cross-referenced by voice—can be relativized, as in (5-a). The object *pitaka* 'wallet' cannot be relativized, as in (5-b). These patterns are consistent with (1) thus far. However, (5-c) is inconsistent with it: the subject *bantay* 'guard' can also be relativized.

(5) When the verb exhibits applicative morphology, the applied argument or the subject can be relativized

a. ang **tao** na dinukutan ng bantay ng pitaka
NOM person LNK steal.APPL GEN guard GEN wallet
'the person from whom the guard stole the wallet'

b. *pitaka na dinukutan ng bantay ang **tao**
wallet LNK steal.APPL GEN guard NOM person
Intended: wallet that the guard stole from the person

c. bantay na dinukutan ng pitaka ang **tao**
guard LNK steal.APPL GEN wallet NOM person
'guard that stole the wallet from the person'

Before proceeding, Ceña and Nolasco (2011) observed that those who allow the subject to be relativized even when not cross-referenced by voice morphology only do so in head-initial RCs, but not in head-final RCs. They reported the following contrasts in (6). Recall that I added *ang* to unambiguously mark the left edge of the RC. When speakers encounter *ang* followed by a verb, they know that they are entering an RC.

(6) In head-initial RCs, the subject may be relativized even when not cross-referenced

by voice; in head-final RCs, it cannot be relativized

a. ang bantay na niloko ang **tao** sa palengke
NOM guard LNK deceive.PV NOM person DAT market
'the guard that deceived the person at the market'

b. *ang niloko ang **tao** sa palengke na bantay
NOM deceive.PV NOM person DAT market LNK guard
Intended: the guard that deceived the person at the market

Recall that the main goal of this chapter is to re-examine the nature of the interaction between voice morphology and A-bar movement in Tagalog. I have outlined the two competing generalizations of the restriction above, and summarize these in (7).

(7) The two competing generalizations of the extraction restriction

a. Voice-match: the argument that is cross-referenced by voice is the only argument eligible for A-bar movement

b. Voice-mismatch: the argument cross-referenced by voice and for some, subjects are eligible for A-bar movement

I framed this interaction in terms of VOICE (MIS)MATCH to preview the terminologies that I will be using in the experiments that follow.

In this chapter, I am primarily interested with two questions. First, I ask whether non-nominative arguments are extractable—and in particular, non-nominative subjects. In a series of acceptability judgment studies, I ask how robust the received view of the extraction restriction is. Under the voice-match restriction, licit extractions in this experiment will only be the extraction of arguments cross-referenced by voice. Extraction of arguments that are not cross-referenced by voice will be considered illicit extractions. Thus, the MISMATCH PENALTY, the contrast between extracting an argument cross-referenced by voice and extracting a co-argument not cross-referenced by voice morphology should

not vary by voice. Framed differently, voice mismatches should be uniformly rejected irrespective of voice. On the other hand, under the voice-mismatch restriction, the mismatch penalty should be greater in AV than in NAV because of the variation in subject extractions under NAV that has been observed.

Second, if non-nominative subjects are extractable, I ask whether this permissibility is a property shared by all A-bar dependencies or whether this is specific to RCs. This is an empirical question because the studies that have observed that the generalization may in fact not be as voice-sensitive mostly involved RCs.

3.2 Mismatches under non-AV voices "in the wild"

Here I offer evidence consistent with view of the extraction restriction that allows voice mismatches. One of my contributions to the discussion on the nature of the extraction restriction in Tagalog is to provide examples of these voice mismatch extractions "in the wild." The examples below are naturally occurring subject extractions under NAV from various sources online, ranging from newspapers and tabloids, to the Bible.

The examples in (8) involve the extraction of the subject when the verb exhibits PV. The examples in (8-a)–(8-b) involve RCs; those in (8-c)–(8-d), topicalizations; and those in (8-e)–(8-f), *ay*-inversions. The examples in (8-c)–(8-f) suggest that the extraction of the subject under PV is not specific to RCs.

(8) Extraction of the subject under PV

a. dyirap na kinain ang **buwan**
giraffe LNK ate.PV NOM moon
'giraffe that ate the moon'

A title of a children's book sold by Amazon: `https://www.amazon.com/Giraffe-That-Ate-Moon-Childrens/dp/1515014258`

b. lalaki-ng alam ang **kanya-ng hangganan**
man-LNK know.PV NOM 3SG.DAT-LNK limit
'man that knows his limit'

A movie teaser for *Goyo: Ang batang heneral* on YouTube (May 5, 2018):

`https://youtu.be/gT6tX_w2OWs?t=1m15s`[4]

c. Hotshots ginulpi ang **Picanto**
PN beat up.PV NOM PN
'Hotshots [a basketball team], they beat Picanto [another team] in landslide.'

A sports article in PhilStar (January 11, 2018): `https://www.philstar.com/pilipino-star-ngayon/palaro/2018/01/11/1776517/hotshots-ginulpi-ang-picanto`

d. truck inararo ang **tindahan sa Bulacan**
truck plowed.PV NOM store DAT PN
'The truck, it plowed through the store in Bulacan'

ABS-CBN news headline (February 18, 2020): `https://news.abs-cbn.com/news/02/18/20/truck-inararo-ang-tindahan-sa-bulacan-mag-ina-patay-12-sugatan`

e. Ang 7% ng mga kabataan ay kinain ang **inirerekomenda-ng dalawa**
NOM 7% GEN PL youth AY ate.PV NOM recommended.PV-LNK two
o higit pa-ng mga servings ng prutas kada araw
or more even-LNK PL serving GEN fruit each day
'As for 7% of the younger people, they eat the recommended two or more servings of fruits each day.'

A post in Just Be Beauty website, *Prutas: Ang bagong cookie* "Fruits: The new cookie": `https://tl.justbebeauty.com/fruit-new-cookie-5529`

f. Ang mga epidemya-ng ito na may dala-ng mga salot ay pinatay
NOM PL epidemic-LNK DEM LNK EXIST carry-LNK PL plague AY killed.PV
ang **halos 25 milyon-g mga Tsino at iba pa-ng mga Asyano**
NOM almost 25 million-LNK PL Chinese and other also-LNK PL Asian
'As for these epidemics that carried the plague, they killed almost 25 million

[4] I thank Henrison Hsieh for bringing this example to my attention.

Chinese and other Asians.'

The Wikipedia entry on *Salot na itim* (Black Death): `https://tl.wikipedia.org/wiki/Salot_na_Itim`

The examples in (9) involve the extraction of the subject when the verb exhibits applicative morphology. The examples in (9-a)–(9-b) involve RCs; those in (9-c)–(9-d), topicalizations; and that in (9-e)–(9-f), *ay*-inversion. The examples in (9-c)–(9-e) suggest that the extraction of the subject under applicative morphology is not specific to RCs.

(9) Extraction of the subject under applicative morphology

a. lalaki-ng kinunan ng cellphone video ang **babae-ng nagsusukat ng swimwear**
man-LNK take.APPL GEN NOM woman-LNK trying on.AV GEN
'man who took a cellphone video of the woman who was trying on a swimsuit'

GMA News headline (November 4, 2019): `https://www.youtube.com/watch?v=9fWP9RWj1gY`

b. parlorista-ng sinabuyan ng kemikal ang **karelasyon ng nobyo**
hairdresser-LNK splashed.APPL GEN chemical NOM partner GEN boyfriend
'hairdresser that splashed chemical on the lover of his boyfriend'

ABS-CBN TV Patrol news (July 20, 2019): `https://www.youtube.com/watch?v=msbky9rN1xc`

c. Andeng, dinalhan ng pagkain si **Marco**
PN bring.APPL GEN food NOM PN
'Andeng (diminutive for Andrea), she brought food to Marco'

Episode recap of TV show *Sandugo*: `https://www.youtube.com/watch?v=_bQcVlASPDU`

d. Police officer sa Michigan, binilhan ng car seat ang **isa-ng**
police officer DAT PN bought.APPL GEN car seat NOM one-LNK

nangangailan na ina
in-need LNK mother
'Police officer in Michigan, he bought a car seat for one mother who was in-need'

News content from TomoNews Philippines (October 14, 2014): `https://youtu.be/utA7FKGQrCg?t=5s`

e. Ang babae po kasi na ibinigay n'yo sa akin ay binigyan
NOM woman HON PART LNK gave.PV 3PL.GEN DAT 1.SG.DAT AY gave.APPL
ako ng bunga
1SG.NOM GEN fruit
'As for the woman who You [God] gave to me, she gave me a fruit'

Ang Salita ng Dios version of the Tagalog Bible (Genesis 3:12): `https://www.bible.com/bible/1264/GEN.3.6-19.ASND`

f. Ang Komisyon sa Wika-ng Filipino (KWF) ay masusi-ng
NOM Commission DAT language-LNK PN AY throrough-LNK
pinagaral-an ang **mga nagdaa-ng ortograpiya-ng Filipino** …
study-APP NOM past-LNK orthography-LNK PN
'As for the Commission on the Philippine language, they carefully examined the previous Filipino orthography (that is currently being enforced in schools).'

Department of Education on National Orthography (14 August 2013): `https://www.deped.gov.ph/wp-content/uploads/2013/08/DO_s2013_034.pdf`[5]

One caveat is that these extracted subjects must surface with nominative marking—overt or otherwise—despite having genitive marking in their post-verbal position. In topicalizations, case marking on the noun is almost always null, though, when present, it needs to be nominative. de Guzman (1995) makes the same observation. Meanwhile, case marking on the noun is obligatory in *ay*-inversion and it needs to be nominative. If these extracted subject surfaced with the marking in their post-verbal position (i.e., genitive), then the extraction is illicit.

[5] I thank Henrison Hsieh for bringing this example to my attention.

3.3 Experiment 1: Comparing voice (mis)match extractions under AV and PV

This section describes what is, to my knowledge, the first systematic investigation of the extraction restriction across a variety of A-bar dependencies in Tagalog. The main goal of this study is to assess the robustness of the received view of the extraction restriction in Tagalog, provided in (1). In order to do so, I manipulated whether the verb exhibited AV or PV and whether the extracted argument was cross-referenced by voice morphology in three types of dependencies (*wh*-questions, RCs, and *ay*-inverted sentences. To assess the robustness of the voice-match restriction, I scaled their ability to detect the illicitness of extraction where the extracted argument is not cross-referenced by the voice (i.e., voice mismatch extractions) against their ability to detect the licitness of extraction where the extracted argument is cross-referenced by the voice (i.e., voice match extractions). Across the three types of dependencies, Tagalog comprehenders were extremely sensitive to mismatches between voice and the extracted argument when the verb exhibited AV. However, their sensitivity varied when the verb exhibited PV. They were more sensitive in *wh*-questions, and then less sensitive in *ay*-inverted sentences and RCs.

3.3.1 Participants

Sixty-four speakers from the University of the Philippines – Diliman participated in this study. They ranged from 18 to 48 years of age (M = 26, SD = 6.5), and all lived in and around Metro Manila at the time of testing. They received a Starbucks gift certificate valued at 200 Philippine peso for participating in the study.

The recruitment criteria were fairly lax. Participants needed to (i) be 18–40 years of age; (ii) live in and around Metro Manila at time of testing; and (iii) use Tagalog every day. It should be noted that people from all over the Philippines live in Metro Manila and

most have some proficiency of Tagalog. The laxness did lead to the inclusion of participants who were also exposed to or possibly even speak other Philippine languages (e.g., Cebuano, Ilokano, etc.). While some would argue that this introduced significant variance in the design of the study, this inclusion does not pose any serious concerns. Most Filipinos whose native language is not Tagalog are early sequential bilinguals. They first learn the language of the region or province where they live, and upon entry into the educational system (around age 5), they learn Tagalog and English (Galang, 1988, 2001). That they were in Metro Manila at the time of testing ensured the active use of Tagalog in their daily lives. I view this inclusion positively because it reflected the social reality of who uses Tagalog in the Philippines, thereby increasing the study's ecological validity.

3.3.2 Materials

The experiment employed a 2×2 design, crossing whether the verb exhibits AV or PV (VOICE: AV, PV), and whether the displaced argument is cross-referenced by voice or not (MATCH: +MATCH, −MATCH). I created three sets of 32 items involving *wh*-questions, relative clauses, and *ay*-inversion (TYPE: WH, AY, RC). These items were randomized by another set of 32 items that are run-of-the-mill declarative sentences (i.e., they did not involve A-bar dependencies. Each item was distributed evenly via Latin square. Apart from the items involving vanilla declaratives, no additional fillers were included. Table 3.1 schematizes the experimental items.

3.3.3 Procedure

The experiment was a pen-and-paper acceptability judgment study. Participants rated the acceptability of sentences using a 7-point Likert-type scale, with 1 being *hindi mabuti* 'not acceptable' (literal: not good) and 7 being *mabuti* 'acceptable' (literal: good).

Table 3.1: Schematization of the items used in Experiment 1. Items in WH, RC , and AY manipulated VOICE (AV, PV), and MATCH (+, –).

TYPE	VOICE	MATCH						
WHQ	AV	+	Which Agent	ang	Verb.AV	XP	GEN-Patient	…
	AV	–	Which Patient	ang	Verb.AV	XP	NOM-Agent	…
	PV	+	Which Patient	ang	Verb.PV	XP	GEN-Agent	…
	PV	–	Which Agent	ang	Verb.PV	XP	NOM-Patient	…
RC	AV	+	… Agent	na	Verb.AV	XP	GEN-Patient	…
	AV	–	… Patient	na	Verb.AV	XP	NOM-Agent	…
	PV	+	… Patient	na	Verb.PV	XP	GEN-Agent	…
	PV	–	… Agent	na	Verb.PV	XP	NOM-Patient	…
AY	AV	+	NOM-Agent	ay	Verb.AV	XP	GEN-Patient	…
	AV	–	NOM-Patient	ay	Verb.AV	XP	NOM-Agent	…
	PV	+	NOM-Patient	ay	Verb.PV	XP	GEN-Agent	…
	PV	–	NOM-Agent	ay	Verb.PV	XP	NOM-Patient	…

3.3.4 Data analysis

I estimated cumulative link mixed models using the *Ordinal* package (Christensen, 2018) in R (R Core Team, 2018). I used their ratings as the dependent measure, and the fixed effects were VOICE (AV, PV), MATCH (+MATCH, –MATCH), TYPE (WHQ, RC, AY), and their interactions. VOICE and MATCH were sum-coded, such that AV and +MATCH mapped to the negative coefficients. TYPE was coded by two coefficients using Helmert contrasts. The first coefficient contrasted WHQ with AY and RC, where WHQ is mapped onto the positive coefficient. The second coefficient contrasted RC with AY, where RC is mapped onto the positive coefficient.[6] I then included the maximal random effects structure justified by

[6]The first coefficient of TYPE contrasted *wh*-questions from head-initial RCs and *ay*-inverted sentences: is the relationship between the filler and the gap direct or mediated. As discussed in Chapter 2, unlike *wh*-questions in English, argument *wh*-questions in Tagalog use a (pseudo-)clefting strategy. The interrogative phrase serves as the predicate and the remaining material is a headless RC. Thus, the relation between the filler and the gap is not direct; it is mediated via predication between what is inside the headless RC and the pronounced *wh*-predicate. In head-initial RCs and *ay*-inverted sentences, the filler-gap relation is more direct. The second coefficient of TYPE contrasted RCs and *ay*-inverted sentences: does the filler have multiple thematic roles in the clause? In RCs, the filler serves one thematic role in the embedded clause and another thematic role in the main clause. In *ay*-inverted sentences, the filler serves only one thematic role.

the design that allowed the models to converge (Barr, 2013; Barr, Gann, & Pierce, 2011). The random effects structure included random intercepts for participants and items, and VOICE, MATCH, and their interaction, and TYPE as random slopes for participants and items.

3.3.5 Results

In Table 3.2 and Figure 3.1, I report the mean ratings by VOICE, MATCH, and TYPE. The inspection of Figure 3.1 reveals that a voice-match view of the extraction restriction is quite robust for AV irrespective of the dependency. The difference between the blue and gold points is consistently large across dependencies, suggesting that participants were rating voice match extractions high and voice mismatch extractions low. On the other hand, a voice-match view of the extraction restriction does not seem tenable for PV. The difference between the blue and gold points varied across constructions. Participants were the most sensitive to mismatches in *wh*-questions, followed by *ay*-inverted sentences, and then by RCs. In Table 3.3, I provide the summary of the estimated model.

Table 3.2: Descriptive statistics of Experiment 1. Mean ratings (with *S.E.*) by VOICE and MATCH.

		Match		
	Voice	+Match	–Match	Penalty
WHQ	AV	4.8 (.2)	2.1 (.2)	2.7
	PV	5.1 (.2)	3.0 (.2)	2.1
RC	AV	5.0 (.2)	2.2 (.4)	2.8
	PV	4.7 (.2)	4.3 (.2)	.4
AY	AV	5.9 (.1)	2.0 (.2)	3.9
	PV	5.3 (.2)	4.5 (.2)	.8

On average, participants gave higher ratings when the verb exhibited PV (M = 4.5) compared to when it exhibited AV (M = 3.7). They gave higher ratings in voice match extractions (M = 5.1) compared to voice mismatch extractions (M = 3.0). They gave higher

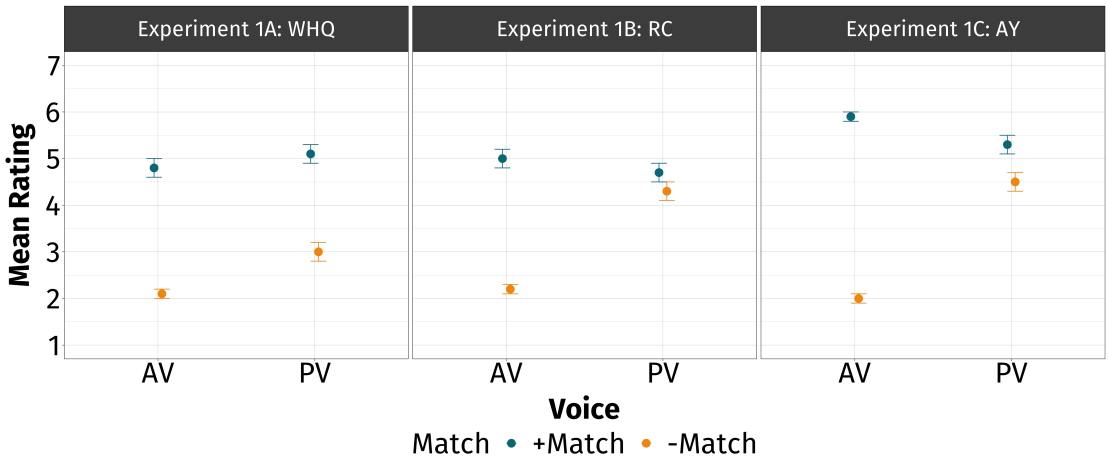


Figure 3.1: Mean rating by VOICE, MATCH, and TYPE in Experiment 1. In blue are the +MATCH-conditions and in gold are the –MATCH-conditions. Standard errors of the mean are also provided. The leftmost panel represents *wh*-questions (experiment 1A), the middle panel represents relative clauses (experiment 1B), and the rightmost panel represents *ay*-inverted sentences (experiment 1C). The robustness of is estimated as the difference between the blue and gold points. The greater the difference, the more robust the restriction.

ratings to relative clauses and *ay*-inverted sentences combined (M = 4.3) compared to *wh*-questions (M = 3.8). They gave higher ratings to *ay*-inverted sentences (M = 4.4) compared to relative clauses (M = 4.1). These effects were qualified by higher-order interactions.

The significant VOICE×MATCH interaction indicated that the difference between -MATCH and +MATCH is greater in AV (difference = 3.1) than it is in PV (difference = 1.09) across TYPE. This suggests that collapsing across the three types of dependencies examined, participants' sensitivity was greater when the verb exhibited AV compared to when it exhibited PV.

The significant MATCH×(WH *vs* RC+AY) interaction indicated that the difference between -MATCH and +MATCH is greater in WHQ (difference = 2.44) than in AY and RC combined (difference = 1.92). This suggests that collapsing across voice, the participants' sensitivity was greater in *wh*-questions than in relative clauses and *ay*-inverted sentences.

The significant MATCH×(RC *vs* AY) interaction indicated that the difference between

Table 3.3: Summary of cumulative link mixed models in Experiment 1. The models included acceptability ratings as the dependent measure, and sum-coded VOICE, MATCH, Helmert-coded TYPE (WH *vs* RC+AY as the first coefficient and RC *vs* AY as the second coefficient), and their interactions into the models.

```
clmm(Rating~Voice * Match * Type + (1 + Voice * Match + Type | Subject)
                + (1 + Voice * Match + Type | Item))
```

	Estimate	*SE*	z
VOICE	1.12	.20	-5.57**
MATCH	-2.88	.33	8.69**
WH *vs* RC+AY	-.48	.19	2.61*
RC *vs* AY	-.58	.24	2.43*
VOICE×MATCH	2.68	.46	-5.85**
VOICE×(WH *vs* RC+AY)	.13	.28	-.48
VOICE×(RC *vs* AY)	-.10	.41	.25
MATCH×(WH *vs* RC+AY)	-1.00	.2 8	3.55**
MATCH×(RC *vs* AY)	1.55	.41	-3.81**
VOICE×MATCH×(WH *vs* RC+AY)	-2.17	.51	4.22**
VOICE×MATCH×(RC *vs* AY)	.04	.79	-.05

*$p < .05$, **$p < .001$

-MATCH and +MATCH is greater in AY (difference = 2.29) than in RC (difference = 1.55). This suggests that collapsing across voice, the participants' sensitivity was greater in *ay*-inverted sentences than in relative clauses.

Finally, the significant VOICE×MATCH×(WH *vs* RC+AY) interaction qualified the two-way interactions found earlier. In other words, the participants' sensitivity in AV was greater than their sensitivity in PV when they encountered relative clauses and *ay*-inverted sentences (difference of difference = 2.72) compared to when they encountered *wh*-questions (difference of difference = .59).

3.3.6 Discussion

To reiterate, I found that across the three types of dependencies, Tagalog comprehenders were extremely sensitive to mismatches between voice and the extracted argument

when the verb exhibited AV. However, their sensitivity varied when the verb exhibited PV. They were more sensitive in *wh*-questions, and then less sensitive in *ay*-inverted sentences and relative clauses.

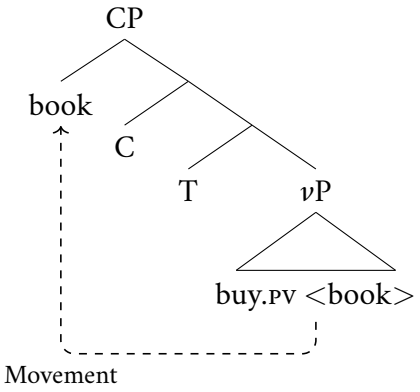
Given what we know about these dependencies, these results are both surprising and unsurprising. It is surprising syntactically and semantically that relative clauses patterned more like *ay*-inverted sentences for two reasons. First, it is widely known in the Austronesian literature that argument *wh*-questions and RCs are structurally similar (Potsdam, 2009). In Tagalog, argument *wh*-questions are pseudo-clefts in which the *wh*-phrase is a predicate nominal and the rest of the material is inside a headless RC (Aldridge, 2002, 2013; Kroeger, 1993). For example, the Tagalog equivalent of the question *Which wine did the woman drink?* would literally be *The one that the woman drank is which wine?* In other words, Tagalog argument *wh*-questions are essentially RCs. Second, the dependency in RCs are thematically more complex than the dependency in *ay*-inverted sentences. The extracted nominal in RCs serves two thematic roles: one in the embedded clause and another one in the main clause. Meanwhile, it only serves one role in *ay*-inverted sentences.

It is also equally unsurprising that RCs patterned more like *ay*-inverted sentences. The dependency between the extracted nominal and the gap is a direct A-bar chain in RCs, following a head raising analysis proposed by Law (2016) and Aldridge (2017) to derive head-initial RCs. This dependency is also a direct A-bar chain in *ay*-inverted sentences. Following the analysis of Sabbagh (2014), I assume that the fronted nominal sits in some specifier position situated in between T and C—either that of a separate functional projection or that of TP. These are schematized in (10) and (11) below.

(10) Adapted from Aldridge (2017)

a. ang **libro**-ng binili ng babae
NOM book-LNK bought.PV GEN woman
the book that the woman bought'

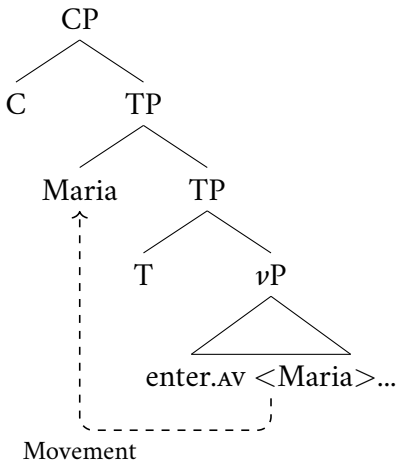
b.



(11) Adapted from Sabbagh (2014)

a. Si Maria ay lumulusong sa bayan upang magtinda
NOM PN AY enter.AV DAT town in.order sell.AV
'Maria goes to town in order to sell things.'

b.

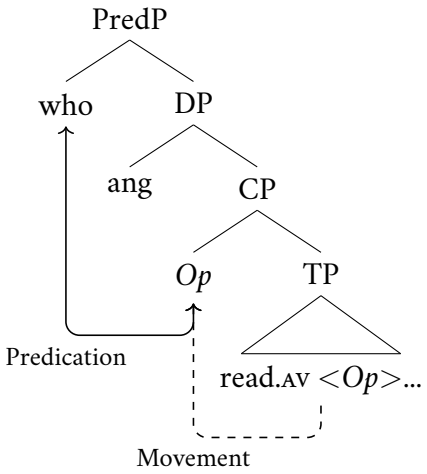


On the other hand, owing to the pseudo-cleft structure of argument *wh*-questions, the dependency between the *wh*-word functioning as a predicate nominal and the gap inside the headless RC is an indirect one and is mediated by a predication relation (Aldridge, 2002; Kroeger, 1993). This indirectness is schematized below in (12).

(12) Adapted from Aldridge (2002)

a. **Sino** ang bumabasa ng libro?
who NOM read.AV GEN book
'Who is reading a book?'

b.



This similarity can be reframed in terms of what is moving. In *ay*-inverted sentence and RCs, following a head-raising analysis, what is moving is overt, as seen in (10) and (11). In argument *wh*-questions, what is moving is a null operator, as seen in (12).

More central to the purposes of this chapter is that these findings cast some doubt on the received generalization of the extraction restriction. These findings are more in line with the descriptions of Ceña and Nolasco (2011, 2012): extractions of the subject are licit under PV for some speakers. These results raise several questions: Are these findings replicable? If they are, do they extend to the other voices in the language (i.e., those with applicative morphology)? How can the asymmetry in the extraction restriction be accounted for? How can the more fine-grained differences among different types of A-bar dependencies be accounted for? What does this tell us about the landscape of A-bar extraction in general?

To answer the first two questions, I conducted another set of judgment studies that seek to replicate and extend these findings. I present in the next section the studies that compare participants' extraction behavior under AV and PV, and under applicatives. While the remaining questions are equally important, this dissertation will only focus on developing an answer for the first two. I establish the necessary groundwork for the third question and present one way the results could be reframed more generally. However, I

leave it to future work to develop a fleshed out account that answers the other questions.

3.4 Experiments 2: Comparing voice (mis)match extractions under AV and PV again

This section presents a second series of acceptability judgment studies, designed to replicate the findings of Experiment 1. Experiment 2 once again compared the extraction behavior of Tagalog speakers when the verb exhibits AV and PV. Another series of judgment studies, presented as Experiment 3 in Section 3.5, were run simultaneously with Experiment 2. Experiment 3 compared the extraction behavior of Tagalog speakers when the verb exhibits applicative morphology.

I investigated their extraction behavior in two contexts, *wh*-questions and head-initial RCs. I chose these two dependencies based on their varying level of permissiveness to mismatch extractions. Experiment 1 showed *wh*-questions to be more compliant to a voice match view of the restriction, and RCs to be less compliant and more consistent with the voice mismatch view. Thus, these two dependencies would give each view of the restriction their best chance to succeed, that is, where they can optimally exhibit their expected effects.

The findings of Experiment 1 were replicated. Participants also gave voice match extractions high ratings; mismatch extractions under AV, low ratings; and mismatch extractions under PV, intermediate ratings.

3.4.1 Participants and procedure

I used the same recruitment criteria as in Experiment 1. Thirty speakers from the University of the Philippines–Diliman and surrounding communities participated in the study. They ranged from 18 to 42 years of age (M = 25, SD = 5), and all lived in and around

Metro Manila at the time of testing. They received either a gift certificate from Bo's (a local coffee shop) or Fully Booked (a local bookstore) valued at 300 Philippine Pesos, for participating in the study.

The judgment studies reported here and in Section 3.5 were run in one session and were administered online via Ibex (Drummond, 2016). Participants were asked to rate the acceptability of sentences using a 7-point scale, with 1 being *talagang hindi katanggap-tanggap* 'really unacceptable' and 7 being *talagang katanggap-tanggap* 'really acceptable.' They were explicitly instructed that a sentence is *katanggap-tanggap* if (i) the order of the elements in the sentence makes sense; (ii) as a speaker of Tagalog, they can imagine themselves saying or writing this; and (iii) they can imagine other Tagalog speakers saying or writing this, as well. On average, each session took about 45 minutes to complete.

3.4.2 Materials

The experiment employed a 2×2 factorial design, crossing whether the verb exhibited AV or PV (VOICE: AV, PV), and whether the extracted argument is cross-referenced by voice or not (MATCH: MATCH, MISMATCH). This is the same design as Experiment 1. I created two sets of 24 items, one with *wh*-questions and one with relative clauses (TYPE: WH, RC), with each item distributed evenly across 4 lists via Latin square design. These were randomized by the two sets of 24 items used in experiment 3, as well as 96 distractor items. Table 3.4 provides sample experimental items for each TYPE.

Half of the distractor items, 2 sets of 24 items, investigated two types of islands, *wh*-islands and complex NP islands. One set compared extraction out of declarative complement clauses (NON-ISLAND) and out of interrogative complement clauses (ISLAND). The other set compared extraction out of complement clauses (NON-ISLAND) and out of complements of gossip-type nominals (ISLAND). These items involved only voice match extractions, and become relevant for the quantitative models estimated in Section 3.6.

Table 3.4: Sample experimental items used in Experiment 2, manipulating VOICE (AV, PV), and MATCH (+, –).

TYPE	VOICE	MATCH						
			Tinanong	ni Pedro	kung …			
			asked.PV	GEN PN	COMP…			
			Pedro asked…					
WHQ	AV	+	sino	ang	kumakanta	ng himno	para sa madre	gabi-gabi.
			who	NOM	sing.AV	GEN hymn	for DAT nun	every night.
			…who sings a hymn for the nun every night.					
	AV	–	ano	ang	kumakanta	ang sakristan	para sa madre	gabi-gabi.
			what	NOM	sing.AV	NOM altar boy	for DAT nun	every night.
			…what the altar boy sings for the nun every night.					
	PV	+	ano	ang	kinakanta	ng sakristan	para sa madre	gabi-gabi.
			what	NOM	sing.PV	GEN altar boy	for DAT nun	every night.
			…what the altar boy sings for the nun every night.					
	PV	–	sino	ang	kinakanta	ang himno	para sa madre	gabi-gabi.
			who	NOM	sing.PV	NOM hymn	for DAT nun	every night.
			…who sings the hymn for the nun every night.					

TYPE	VOICE	MATCH					
			Medyo baduy	ang …			
			somewhat tacky	NOM			
			The …is somewhat tacky.				
RC	AV	+	manliligaw	na	kumakanta	ng kundiman	para sa nililigawan.
			suitor	LNK	sing.AV	GEN love song	for DAT wooee.
			…suitor who sings a love song for the person they are wooing …				
	AV	–	kundiman	na	kumakanta	ang manliligaw	para sa nililigawan.
			love song	LNK	sing.AV	NOM suitor	for DAT wooee.
			…love song that the suitor sings for the person they are wooing …				
	PV	+	kundiman	na	kinakanta	ng manliligaw	para sa nililigawan.
			love song	LNK	sing.PV	GEN suitor	for DAT wooee.
			…love song that the suitor sings for the person they are wooing …				
	PV	–	manliligaw	na	kinakanta	ang kundiman	para sa nililigawan.
			suitor	LNK	sing.PV	NOM love song	for DAT wooee.
			…suitor who sings the love song for the person they are wooing …				

3.4.3 Data analysis

I estimated cumulative link mixed models using the package *Ordinal* (Christensen, 2018) in R (R Core Team, 2018). We used their ratings as the dependent measure, and the fixed effects were Voice (av, pv), Match (+match, −match), Type (whq, rc), and their interactions. Voice, Match and Type were sum-coded, such that av, +match and whq mapped to the negative coefficients. I then included the maximal random effects structure justified by the design that allowed the models to converge (Barr et al., 2011; Barr, 2013).

3.4.4 Results

In Table 3.6 and Figure 3.2, I report the mean ratings by voice, match, and type. In Table 3.7, I provide the summary of the estimated model.

Table 3.6: Mean ratings (with standard errors of the mean) by Voice (AV, PV) and Match (+Match, −Match). Penalty score is derived by subtracting −Match from +Match

		Match		
	Voice	+Match	−Match	Penalty
whq	AV	6.4 (.1)	2.1 (.1)	4.3
	PV	6.3 (.1)	3.0 (.1)	3.3
rc	AV	5.4 (.2)	2.2 (.1)	3.2
	PV	6.0 (.1)	4.3 (.2)	1.7

There were two main effects: that of Match and Voice. On average, participants gave higher ratings to +match extractions (M = 6.0) than −match extractions (M = 2.9). They gave higher ratings to extractions with pv (M = 4.9) than those with av (M = 4.0). They gave comparable ratings to rc (M = 4.5) and whq (M = 4.4). These effects were qualified by higher-order interactions.

All three of the two-way interactions were statistically significant. The significant Match×Voice interaction indicated that the mismatch penalty is greater in av (difference = 3.7) than it is in pv (difference = 2.4) across Type. This suggests that collapsing the type

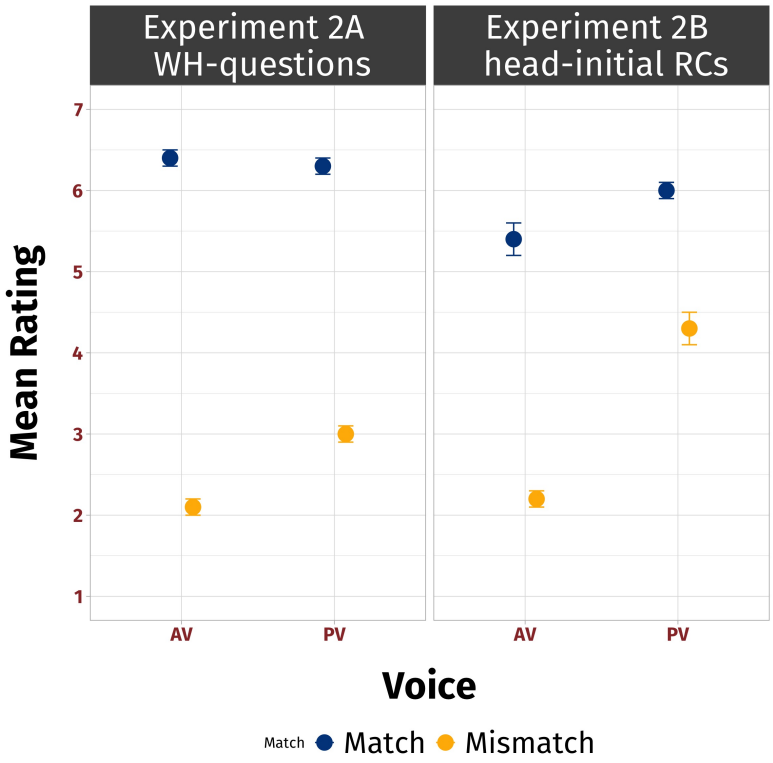


Figure 3.2: Mean rating by VOICE, MATCH, and TYPE in Experiment 2. In blue are the +MATCH-conditions and in gold are the –MATCH-conditions. Standard errors of the mean are also provided. The left panel represents *wh*-questions, and the right panel represents represents relative clauses. The robustness of is estimated as the difference between the blue and gold points. The greater the difference, the more robust the restriction.

of A-bar dependency, the participants' mismatch extractions incurred a greater penalty in AV than in PV. This is consistent with what I found in Experiment 1. I take this as evidence that the extraction restriction is not clean-cut, contra the received view of the extraction restriction.

The significant MATCH×TYPE interaction indicated that the mismatch penalty in WHQ (difference = 3.8) is greater than that in RC (difference = 2.4). This suggests that collapsing across VOICE, the participants' mismatch extractions incurred greater penalty in *wh*-questions than in relative clauses.

The significant VOICE×TYPE interaction indicated that the difference between AV and PV is greater in RC (difference = 1.3) than in WHQ (difference = .4). This suggests that

Table 3.7: Summary of the cumulative link mixed models estimated in Experiment 2

```
clmm(Rating~Voice * Match * Type + (1 + Voice * Match + Type | Subject)
                + (1 + Voice * Match + Type | Item))
```

	Estimate	*SE*	z
MATCH	-4.70	.36	-13.21**
VOICE	1.13	.20	5.63**
TYPE	-.13	.24	-.52
MATCH×VOICE	1.77	.52	3.41**
MATCH×TYPE	2.28	.31	7.44**
VOICE×TYPE	1.44	.32	4.48**
MATCH×VOICE×TYPE	-.04	.88	-.04

*$p < .05$, **$p < .001$

collapsing across MATCH, the participants' ratings for PV were higher than their ratings for AV in relative clauses than in *wh*-questions.

The three-way interaction was not statistically significant. This suggests that in the current study, I do not have evidence to claim that the asymmetry differed by A-bar type.

3.4.5 Discussion

The main goal of Experiment 2 was to see whether the findings of Experiment 1 were replicable. I found that voice mismatch extractions incurred a greater penalty in AV than in PV. This is consistent with the results of Experiment 1. However, unlike Experiment 1, there was no evidence that RCs exhibited greater asymmetry than *wh*-questions. It could be that what I found in Experiment 1 was a fluke and that there is no real difference in how permissive *wh*-questions and relative clauses are with respect to voice mismatch extractions. However, there is reason to believe that there is a real difference there. This goes against what I have found in my own fieldwork. For those speakers who consistently allow subject extractions under PV, they have consistently done so with RCs. With *wh*-questions, their behavior was more variable. Sometimes the same speaker would accept

the same sentence one day and reject it another day. One generalization that has been *relatively* consistent is that voice mismatches seem to fare better in *d*-linked questions than in bare interrogatives. Thus, I am committed to the differences found in Experiment 1 based on my own work with speakers. Then the question becomes why these differences were not observed in Experiment 2.

I conjecture that the underlying cause for the difference is the slight differences in the experimental items used. Recall that in Experiment 1, I used matrix *d*-linked *wh*-questions, while in Experiment 2, I used embedded bare *wh*-questions. One obvious way that the items differ is that one used *d*-linked questions and one used bare interrogatives. If my observation that voice mismatches seem to fare better *d*-linked questions than in bare interrogatives is correct, these results become even more curious. If bare interrogatives were more resistant to voice mismatches, then we should even greater WHQ-RC difference in Experiment 2 than in Experiment 1. A related construal of how this lack of effect in Experiment 2 could be artifactual is that it might have nothing to do with the difference between *d*-linked and bare interrogatives but is more about the difference between embedded and matrix questions. I have not done any systematic investigation or fieldwork to have any concrete predictions. Another related construal is that it could be related to the difference in both the type of interrogative used and the embedding context. I leave these questions open for now.

To sum up, the general results of Experiment 1 were replicated. Experiment 2 showed that PV incurred greater mismatch penalties compared to AV. However, I found no evidence for varying levels of asymmetry by A-bar type.

3.5 Experiments 3: Comparing voice (mis)match extractions under applicatives

This section presents a third series of acceptability judgment studies, designed to extend the findings of Experiments 1 and 2 to the other voices in the language. Recall that Experiments 2 and 3 were run simultaneously. In short, I found that the voice-mismatch view of the extraction restriction also extends to the other voices in the language. Participants gave match extractions high ratings; extractions of the object under applicative, low ratings; and extractions of the subject under applicative, intermediate ratings.

3.5.1 Participants and procedure

The same 30 speakers participated in the study. The same procedure was used.

3.5.2 Materials

The experiment employed a 2×2 factorial design, crossing whether the extracted argument is cross-referenced by voice or not (MATCH: +MATCH, −MATCH) and whether extracted argument is the subject/agent or not (AGEX: YES, NO). The design of Experiment 3 is very different from that of Experiments 1 and 2, but the basic idea of calculating mismatch penalties remain the same. I make explicit the comparisons that I am making first in prose and then I summarize them in Table 3.8.

When a verb has applicative morphology, there are three logical possibilities for which argument can be extracted: (i) the applied object, which is cross-referenced by voice morphology; (ii) the genitive-marked subject; and (ii) the genitive-marked object. Because I am interested in seeing whether subjects can be extracted under applicatives, I need a baseline for extractions under applicatives, one where the extracted argument is cross-referenced by applicative morphology. Thus, one of the conditions of the experiment

looked at the extraction of applied objects under applicatives. I also need a baseline for subject extraction, one where the extracted subject is cross-referenced by voice morphology. Thus, one of the conditions of the experiments looked at subject extraction under AV. Now that I have these baselines, I can compare them to the actual condition of interest, that is, to subject extraction under applicatives. The remaining condition is when it is not voice-matched and it is not a subject extraction. Extraction of the object, the last logical possibility for what argument can be extracted under applicatives, fits the bill.

Table 3.8: Overview of how the experimental conditions in Experiment 3 mapped onto the types of extractions

	Match	
AgEx	+match	−match
yes	Agent extraction under AV	Agent extraction under applicative
no	Applied object extraction under applicative	Patient extraction under applicative

I created two sets of 24 items, one with *wh*-questions and one with relative clauses (Type: wh, rc), with each item distributed evenly across 4 lists via Latin square design. These were randomized by the two sets of 24 items used in Experiment 2, as well as 96 distractor items. Table 3.9 provides sample experimental items for each Type.

3.5.3 Data analysis

I estimated cumulative link mixed models using the package *Ordinal* (Christensen, 2018) in R (R Core Team, 2018). I used their ratings as the dependent measure, and the fixed effects were Match (+match, −match), AgEx (yes, no), Type (whq, rc), and their interactions. Match, AgEx and Type were sum-coded, such that +match, yes and whq mapped to the negative coefficients. I then included the maximal random effects structure justified by the design that allowed the models to converge.

Table 3.9: Sample experimental items used in Experiment 3, manipulating MATCH (+, –) and AGEX (YES, NO) .

TYPE	MATCH	AGEX	Sinabi	ni Maria	kung …			
			said.PV	GEN PN	COMP…			
			Maria said…					
	+	YES	sino	ang	bumibili	ng mangga	para sa ulila	buwan-buwan.
			who	NOM	buy.AV	GEN mango	for DAT ophan	every month.
			…who buys mango for the orphan every month.					
	+	NO	sino	ang	binibilhan	ng magsasaka	ng mangga	buwan-buwan.
			who	NOM	buy.APPL	GEN farmer	GEN mango	every month.
WHQ			…who the farmer buys mango for every month.					
	–	YES	sino	ang	binibilhan	ang ulila	ng mangga	buwan-buwan.
			who	NOM	buy.APPL	NOM orphan	GEN mango	every month.
			…who buys mango for the orphan every month.					
	–	NO	ano	ang	binibilhan	ng magsasaka	ang ulila	buwan-buwan.
			what	NOM	buy.APPL	GEN farmer	NOM orphan	every month.
			…what the farmer buys for the orphan every month.					

TYPE	MATCH	AGEX					
			Bongga	ang …			
			extravagant	NOM			
			The …is extravagant.				
	+	YES	biyenan	na	bumili	ng house and lot	para sa bagong kasal.
			in-law	LNK	bought.AV	GEN house-and-lot	for DAT newly wed.
			…in-law who bought a house-and-lot for the newly wed …				
	+	NO	bagong kasal	na	binilhan	ng biyenan	ng house and lot.
RC			newly wed	LNK	bought.APPL	GEN in-law	GEN house and lot.
			…newly weds that the in-law bought a house-and-lot for …				
	–	YES	biyenan	na	binilhan	ang bagong kasal	ng house and lot.
			in-law	LNK	bought.APPL	NOM newly wed	GEN house and lot.
			…in-law that bought a house-and-lot for the newly wed …				
	–	NO	house-and-lot	na	binilhan	ng biyenan	ang bagong kasal.
			house-and-lot	LNK	bought.APPL	GEN in-law	NOM newly wed.
			…house-and-lot that the in-law bought for the newly wed …				

3.5.4 Results

In Table 3.11 and in Figure 3.3, I report the mean ratings by MATCH, AGEX, and TYPE.

In Table 3.12, I provide the summary of the estimated model.

Table 3.11: Mean ratings (with standard errors of the mean) by MATCH (+MATCH, −MATCH) and AGEX (YES, NO). Penalty score is derived by subtracting −MATCH from +MATCH

		Match		
	AGEX	+Match	−Match	Penalty
WHQ	YES	6.2 (.1)	3.6 (.2)	2.6
	NO	5.9 (.1)	2.2 (.1)	3.7
RC	YES	5.2 (.2)	4.2 (.1)	1.0
	NO	4.9 (.2)	2.5 (.1)	2.4

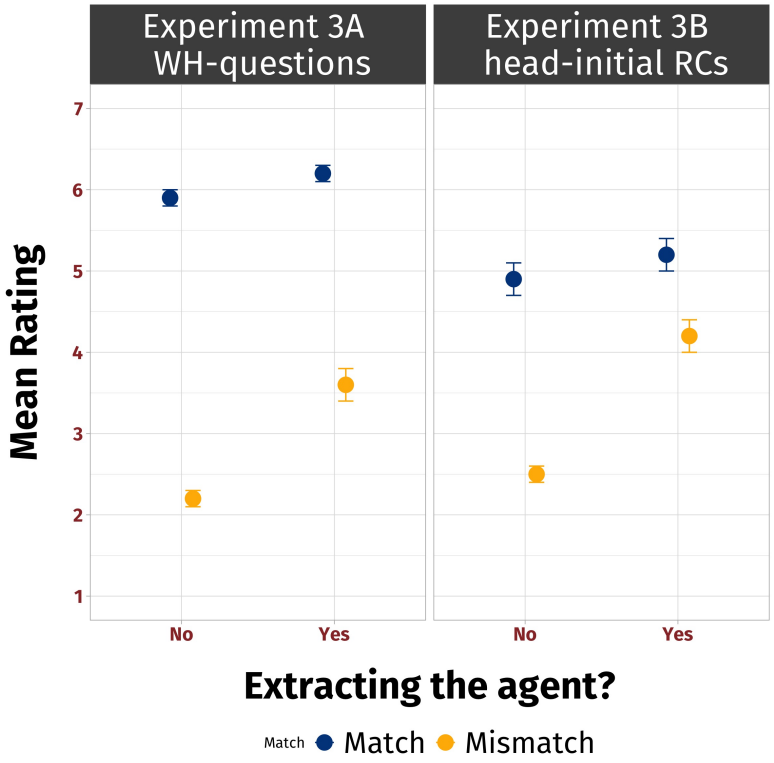


Figure 3.3: Mean rating by AGEX, MATCH, and TYPE in Experiment 3. In blue are the +MATCH-conditions and in gold are the −MATCH-conditions. Standard errors of the mean are also provided. The left panel represents *wh*-questions, and the right panel represents represents relative clauses.

Table 3.12: Summary of the cumulative link mixed models estimated in Experiment 3. The models included acceptability ratings as the dependent measure, and sum-coded MATCH and AGEX, and their interactions into the models.

```
clmm(Rating~Match * AgEx * Type + (1 + Match * AgEx + Type | Subject)
                + (1 + Match * AgEx + Type | Item))
```

	Estimate	*SE*	z
MATCH	-3.29	.30	-11.02**
AGEX	-1.38	.22	-6.14**
TYPE	-.42	.26	-1.59
MATCH×AGEX	-1.31	.46	-2.87*
MATCH×TYPE	2.01	.31	6.51**
AGEX×TYPE	-.15	.35	-.42
MATCH×AGEX×TYPE	-.52	.53	-.99

*$p < .05$, **$p < .001$

There were two main effects: that of MATCH and AGEX. On average, participants gave higher ratings to +MATCH extractions (M = 5.6) than −MATCH extractions (M = 3.1). They gave higher ratings to subject/agent extractions M = 4.8) than non-subject/agent extractions AV (M = 3.9). They gave comparable ratings to RC (M = 4.2) and WHQ (M = 4.5). These effects were qualified by higher-order interactions.

Two of the three two-way interactions were statistically significant. The significant MATCH×AGEX interaction indicated that the mismatch penalty is greater when the extracted argument is not the agent (NO; difference = 3.1) than when it is (YES; difference = 1.8), collapsing across TYPE. This suggests that collapsing the type of A-bar dependencies, the participants' mismatch extractions incurred a greater penalty in non-agent extraction than in agent extraction. This echoes what Experiment 1 found. I take this as evidence that the extraction restriction is not clean-cut, contra the received view.

The significant MATCH×TYPE interaction indicated that the mismatch penalty is greater in WHQ (difference = 3.1) than it is in RC (difference = 1.8) across TYPE. This suggests that collapsing the type of A-bar dependency, the participants' mismatch extractions incurred

a greater penalty in *wh*-questions than in relative clauses.

The three-way interaction was not statistically significant. As in Experiment 2, there is no evidence that the asymmetry differed by A-bar type.

3.5.5 Discussion

In Experiment 3, I found that voice mismatch extractions under applicatives incurred a greater penalty when the extracted argument was the object, compared to when it was the subject. Much like Experiment 2, there was no evidence that relative clauses exhibited greater asymmetry than *wh*-questions. Since this is the first experimental investigation of the extraction behavior of Tagalog speakers under applicatives, it is difficult to ascertain why I did not observe varying levels of asymmetry by A-bar type. One reason could be that the varying degrees of asymmetry that Experiment 1 found in AV and PV are leveled in applicatives. Or it may be artifactual, as discussed in Experiment 2. Recall that Experiment 1 used matrix *d*-linked *wh*-questions, while Experiments 2 and 3 used embedded bare *wh*-questions.

To sum up, the general results of Experiment 1 extended to the other voices in the language. Experiment 3 showed that the mismatch penalty was greater when the extracted argument was the object, compared to when it was the subject, under applicatives.

Before anyone can start developing an account for the extraction restriction that allows voice mismatches, it is crucial that we understand the nature of the ratings distribution that underlies the intermediate mean ratings of subject extractions under NAV. Intermediate mean ratings are compatible with (at least) two types of response distributions. In one scenario, participants classified these extractions categorically, and the intermediate ratings stemmed from a mixture of acceptable and unacceptable responses. In another scenario, participants perceived these extractions as somewhere between fully acceptable and unacceptable. The intermediate ratings reflected an overall upward shift in

acceptability relative to uncontroversially unacceptable extractions or an overall downward shift in acceptability relative to uncontroversially acceptable extractions. In the next section, I develop quantitative models to estimate the participants' judgment process to better understand the underlying distributions of their ratings.

3.6 A model of the participants' judgment process

This section develops quantitative models using Dillon et al. (2017)'s framework to estimate the underlying distribution of the participants' ratings of subject extractions under PV reported in Experiments 1–3. In general, intermediate ratings are compatible with (at least) two types of distributions:

- A mixture of categorical responses;
- An overall shift in acceptability.

These two possible scenarios are illustrated in Figure 3.4. Following Dillon et al. (2017), I refer to the first scenario as having a DISCRETE distribution, and the second scenario, as having a GRADIENT distribution.

I found that the discrete models provided a closer fit to the empirical response distribution compared to the gradient model. I take these results as evidence that in any given trial, participants either accepted or rejected subject extractions under NAV. These results, taken together with the descriptions made by Ceña and Nolasco (2011, 2012), are consistent with there being two types of extraction grammars in Tagalog: one that is consistent with the voice match restriction in disallowing agent extractions under NAV, and one that is consistent with a restriction that allows voice mismatch extractions under NAV, with the proviso that the extracted argument is the subject.

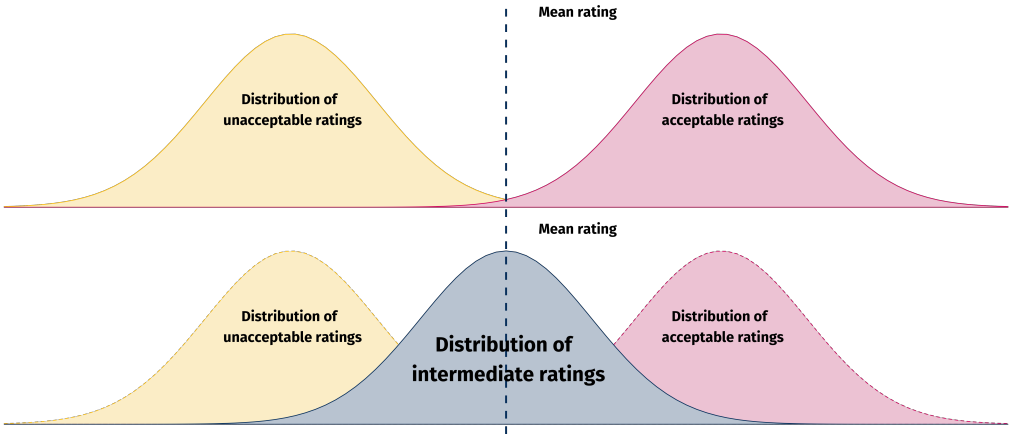


Figure 3.4: Two potential underlying distributions of intermediate ratings. *Top panel:* A discrete model of the participants' judgment process, where intermediate ratings are a result of a mixture of categorical responses. *Bottom panel:* A gradient model of participants' judgment process, where intermediate ratings reflect an overall shift in acceptability. The yellow curves represent the unacceptable distribution. The red curves represent the acceptable distribution. The blue curve represents the distribution of intermediate ratings under a gradient model

3.6.1 Method

To determine which model best captured the response behavior of the participants, three types of distributions are needed as ingredients for the simulations: (i) a test distribution, which is the distribution of the linguistic object that we want to model; (ii) an acceptable reference distribution that gives us an estimate of the sampling space when a linguistic object is uncontroversially acceptable; and (iii) an unacceptable reference distribution that gives us an estimate of the sampling space when a linguistic object is uncontroversially unacceptable.

For experiment 1, I combined the ratings for voice match extractions (i.e., subject extractions under AV and object extractions under PV) and used them as the grammatical reference distribution (*A*) because these types of extractions can give us an estimate of the sampling space when an extraction is uncontroversially acceptable. I then used the ratings for object extractions under AV as the unacceptable reference distribution (*U*)

because this type of extraction can give us an estimate of the sampling space when an extraction is uncontroversially unacceptable. I followed closely how Dillon et al. generated the distributional predictions of these response models.

For experiments 2 and 3, I used the ratings for extractions out of complement clauses (i.e., non-islands) as the grammatical reference distribution. I then used the ratings for extractions out of *wh*-islands and complex NP-islands (i.e., gossip-type NPs) as the ungrammatical reference distributions. These ratings were obtained in tandem with the ratings reported in Experiments 2 and 3. These items were half of the filler items. Below in Table 3.13 I provide a summary of their mean ratings. I again followed closely how Dillon et al. generated the distributional predictions of these response models.

Table 3.13: Mean ratings (with standard errors of the mean) by VOICE (AV, PV) and ISLAND-EXTRACTION (YES, NO).

			Voice	
	ISLAND?	Comparison	AV	PV
CNPC	NO	Clausal comp.	5.5 (.1)	4.6 (.2)
	YES	Complex NP	2.4 (.1)	2.3 (.1)
WH-islands	NO	Declarative CP	5.5 (.2)	5.1 (.2)
	YES	Interrogative CP	3.1 (.1)	2.6 (.1)

Under a discrete model of the participants' response behavior, I assumed that every rating for subject-extraction under PV/applicative is an independent draw from either the acceptable ratings distribution *A* or the unacceptable ratings distribution *U*. The sampling probability is controlled by the mixing parameter π, defined here as the probability of drawing from *A*. A sample judgment *j* is computed using the equation in (13):

(13)
$$\begin{cases} j \sim A & \pi \\ j \sim U & 1 - \pi \end{cases}$$

Under a gradient model, I assumed that every rating for subject-extraction under PV/applicative is an independent draw from the samples *a* and *u* from the acceptable ratings distribu-

tion *A* or the unacceptable ratings distribution *U*, respectively. A sample judgment *j* is a weighted combination from the reference distributions, where π is the relative weight given to samples drawn from *A*, and is computed using the equation in (14):

$$
\begin{aligned}
& a \sim A \\
& u \sim U \\
& j = a\pi + u(1 - \pi)
\end{aligned}
\tag{14}
$$

As a metric for evaluating how well a given model fits the empirical distribution, I calculated the difference in the models' Bayesian Information Criterion (BIC).

3.6.2 Judgment process in Experiment 1

In a nutshell, I found that across dependency types, the discrete models provide a closer fit to the empirical response distribution compared to the gradient model.

I pooled the observations across participants within conditions. I fitted discrete and gradient models to subject extractions under PV for each TYPE. I ran 1000 Monte Carlo simulations of the models, and I report the summary statistics of the two response models in Table 3.14. Note that the values reported here are calculated over these simulations.

Even though both models have very different underlying distributions, they achieved a very close fit between the predicted mean and the empirical mean for subject extractions under PV for each TYPE. Metrics to compare the models' goodness of fit indicate an advantage for the discrete model over the gradient model. Provided in Table 3.14 is the difference in the models' Bayesian Information Criterion (BIC). Negative BIC difference scores indicate an advantage for the discrete model over the gradient model. On the other hand, positive BIC difference scores indicate an advantage for the gradient model over the discrete model. Of the 1000 simulations that I ran, all 1000 BIC difference scores were negative for WHQ and RC; and 998 BIC difference scores were negative for AY.

Table 3.14: Summary of the response models for Experiment 1. Values represent the average over 1000 Monte Carlo simulations. The parameter π and the predicted mean under each model are provided. Square brackets represent the range of values observed across Monte Carlo trials

		WHQ	RC	AY
	Empirical mean	3.0	4.5	4.3
DISCRETE	π	.7	.3	.2
	Predicted mean	3.0	4.5	4.3
GRADIENT	π	.7	.3	.2
	Predicted mean	3.1	4.5	4.3
BIC difference score		-108.1	-52.4	-100.8
$BIC_{Discrete} - BIC_{Gradient}$		[-168.5, -50.5]	[-136.1, 14.8]	[-188.3, -38.6]

Grammatical reference distribution = Subject-extractions under AV and Object-extractions under PV; Ungrammatical reference distribution = Object-extractions under AV

3.6.3 Judgment process in Experiment 2

In a nutshell, I found that across dependency types, the discrete models provide a closer fit to the empirical response distribution compared to the gradient model.

I pooled the observations across participants within conditions. I fitted discrete and gradient models to subject extractions under PV for each TYPE. I ran 1000 Monte Carlo simulations of the models, and I report the summary statistics of the two response models in Table 3.15. Note that the values reported here are calculated over these simulations.

Unlike the first simulation where the predicted means that the models generated were quite spot on, the predicted means generated by the models for Experiment 2's subject extraction under PV were not quite spot on. The discrete model's predicted mean was smaller than the empirical mean, while the gradient model's predicted mean was larger than the empirical mean.. Metrics to compare the models' goodness of fit indicate an advantage for the discrete model over the gradient model. Provided in Table 3.15 is the

Table 3.15: Summary of the response models for Experiment 2. Values represent the average over 1000 Monte Carlo simulations. The parameter π and the predicted mean under each model are provided. Square brackets represent the range of values observed across Monte Carlo trials

		WHQ	RC
	Empirical mean	3.0	4.3
DISCRETE	π	.63	.63
	Predicted mean	3.5	3.5
GRADIENT	π	.22	.75
	Predicted mean	3.99	5.27
BIC difference score		-15.79	-23.72
$BIC_{Discrete} - BIC_{Gradient}$		[-77.21, 76.32]	[-49.87, 23.38]

Grammatical reference distribution = Extraction out complement clauses (non-islands);
Ungrammatical reference distribution = Extraction out of islands (CNPC and *wh*-islands)

difference in the models' BIC. Recall that negative BIC difference scores indicate an advantage for the discrete model over the gradient model, while positive BIC difference scores indicate an advantage for the gradient model over the discrete model. Of the 1000 simulations that I ran, 730 BIC difference scores were negative for WHQ, and 977 BIC difference scores were negative for RC.

3.6.4 Judgment process in Experiment 3

In a nutshell, I found that across dependency types, the discrete models provide a closer fit to the empirical response distribution compared to the gradient model.

I pooled the observations across participants within conditions. I fitted discrete and gradient models to subject extractions under applicatives for each TYPE. I ran 1000 Monte Carlo simulations of the models, and I report the summary statistics of the two response models in Table 3.16. The values reported here are calculated over these simulations.

Table 3.16: Summary of the response models for Experiment 3. Values represent the average over 1000 Monte Carlo simulations. The parameter π and the predicted mean under each model are provided. Square brackets represent the range of values observed across Monte Carlo trials

		WHQ	RC
	Empirical mean	3.6	4.2
DISCRETE	π	.63	.62
	Predicted mean	3.5	3.49
GRADIENT	π	.42	.76
	Predicted mean	4.57	5.23
BIC difference score		-14.92	-36.14
$BIC_{Discrete} - BIC_{Gradient}$		[-61.1, 50.24]	[-72.39, 4.65]

Grammatical reference distribution = Extraction out complement clauses (non-islands);
Ungrammatical reference distribution = Extraction out of islands (CNPC and *wh*-islands)

Unlike the first simulation where the predicted means that the models generated were quite spot on, the predicted means generated by the models for Experiment 3's subject extraction under applicatives were not quite spot on. Like the simulations for Experiment 2, the discrete model's predicted mean was smaller than the empirical mean, while the gradient model's predicted mean was larger than the empirical mean. Metrics to compare the models' goodness of fit indicate an advantage for the discrete model over the gradient model. Provided in Table 3.16 is the difference in BIC. Recall that negative BIC difference scores indicate an advantage for the discrete model over the gradient model, while positive BIC difference scores indicate an advantage for the gradient model over the discrete model. Of the 1000 simulations, 783 BIC difference scores were negative for WHQ, and 998 BIC difference scores were negative for RC.

3.6.5 Discussion

The results from these simulations suggest that speakers were most likely treating subject extractions under PV and under applicatives categorically. That is, speakers were either accepting or rejecting them in any given trial. One interpretation consistent with these findings is that there could be two grammars of extraction operational in Tagalog—one grammar that requires extraction of arguments to be voice-matched; and the other grammar allows for voice-mismatched extractions, so long as it is the subject. Alternatively, speakers may fail to find the correct derivation for these extractions because of an independently attested advantage of agents in comprehension and production in Tagalog (Sauppe, 2016). When these results are taken together with my own fieldwork and previous descriptions about the variation in extraction restriction, I argue that a two-grammar hypothesis is more tenable.

3.7 General discussion and conclusion

To reiterate, this chapter found the following:

- In Experiments 1 and 2, I compared voice-match extractions and voice-mismatch extractions and found that PV incurred greater mismatch penalties than AV.
- In experiment 3, I found that extracting the object incurred greater mismatch penalties than extracting the subject when the verb had applicative morphology.
- Results from the simulations suggest that participants were treating subject extractions under PV and applicatives categorically. That is, in any given trial, they were either rejecting or accepting these types extraction.

These results are consistent with the description by Ceña and Nolasco (2011, 2012) that the extraction restriction is not clean-cut. I took these results as evidence that there could

be two grammars of Tagalog.

The rest of the section is organized as follows. I first try and reformulate these two competing extraction restrictions in terms of grammar competition in learning, in the vein of Han et al. (2007). I then talk about the implications of these findings for the formal analyses of the extraction restriction. Lastly, I discuss the implications of these findings for how voice could be used in real-time.

3.7.1 A reformulation: Two grammars?

Here I present one way to reframe the competing extraction restrictions. One way to interpret the results is that there could be two grammars of extraction operational in Tagalog. Let us consider the two extraction restrictions in terms of grammar competition in learning in the sense of Han et al. (2007). This section is speculative in nature, as no data has been collected yet on how children acquire the extraction restriction.

The main intuition behind grammar competition in learning is that sometimes the primary linguistic data that a child in a language L is exposed to are compatible with at least two hypotheses, H_1 and H_2. When linguistic experience cannot adjudicate between the two, some learners acquire a grammar based on H_1 and others, one based on H_2.

Han et al. (2007) investigated Korean verb-movement as a case study. One of the challenges of learning a verb-final language like Korean is ascertaining where the verb is in the syntactic structure. When learners of Korean—or other verb-final languages—are confronted with this puzzle, their linguistic data are compatible with (at least) two hypotheses: (i) VERB RAISING hypothesis: the verb undergoes head-movement but this movement is oftentimes string vacuous; or (ii) IN SITU hypothesis: the verb does not undergo head-movement and remains in its base-generated position.

One piece of evidence that could potentially adjudicate between the two involves the interaction between negation and quantifiers. These types of scope interactions are rare,

however, in a Korean child's input (Han et al., 2007, p. 2). With little or no evidence favoring one over the other, child learners develop different grammars.

Han et al. investigated scope interactions in child- and adult-Korean using a truth-value judgment task. Their results show that "only about half of the Korean speakers allow negation to take scope over an object QP" (p. 32). They interpret this as evidence consistent with a split-grammar in the Korean. A subset of the Korean-speaking population has developed a verb-raising grammar, and another subset, an in situ grammar.

I conjecture that the variation in the extraction restriction in Tagalog can also be reframed in terms of grammar competition in learning. In Chapter 2, I surveyed four recent approaches to voice morphology and how they account for voice-match extractions in Tagalog. In a subsection below, I sketch out various modifications to existing proposals in order to account for voice-mismatch extractions. One could think of these modifications as descriptions of what the other Tagalog grammar of extraction could look like. Here, I simply preview where the locus of variation could reside, which largely depends on the approach to voice morphology and voice-match restriction.

Under a multiple specifier approach, one place that could be burdened to be the locus of variation would be the mechanism by which we build our syntactic trees. Implicit in multiple specifier approaches to voice morphology and the voice match restriction is the "no tucking-in" assumption in derivations when the object or the applied object is moved to the outermost specifier of *v*P. This could reflect one grammar of Tagalog. The other grammar could very well have both "no tucking-in" and "tucking-in" derivations available when an argument is moved to the outermost specifier of *v*P.

Another place that could be burdened to be the locus of variation under approaches with multiple specifiers would be how we compute locality in these configurations. Under this view, there could be two types of Tagalog speakers and they differ in how they view multiple specifier configurations with respect to probe-goal relations. One grammar only

sees the outermost specifier, while the other grammar can see both specifiers.

Under an approach that sees voice as agreement with C, we can burden how functional heads bundle features. One grammar could bundle [$u\phi$] and [EPP] together and thus, they probe together and are licensed by the same DP, while the other grammar does not bundle [$u\phi$] and [EPP] and thus, they may probe separately and therefore, can be licensed by different DPs.

3.7.2 Implications for formal analyses of the extraction restriction

In Chapter 2, I provided an overview of different formal implementations of voice match extractions (*i.e.,* extraction of the subject under AV, of the object under PV, and of the applied object under applicative). Irrespective of which proposal is adopted, these types of extractions are standardly analyzed as being derived via A-bar movement. Here I first establish that Tagalog is an island-sensitive language. I then provide data suggesting that the extraction of the subject under PV and applicatives is also island-sensitive. I then sketch out some modifications to these formal analyses.

Tagalog is an island-sensitive language

The seminal work of Ross (1967) showed that there are constraints on movement. That is, there are syntactic configurations where movement is blocked. Here I first show that Tagalog is an island-sensitive language, independent of subject extractions under NAV. Showing that Tagalog is an island-sensitive language, independent of subject extractions under NAV, is a necessary step to argue for a movement-based account because some Austronesian languages have been reported to be island-insensitive—Palauan, as described by Georgopoulos (1985); Isbukun Bunun, as described by H.-h. I. Wu (2016); and Puyuma and Amis, as described by V. Chen and Fukuda (2016). Lastly, I demonstrate that *for those who accept extractions of the subject under NAV,* they also exhibit island-effects.

First, Tagalog is sensitive to *wh*-islands, as shown in (15). In (15-a) and (15-b), *tinapay* 'bread' is in its canonical post-verbal position. The crucial difference between the two is that in (15-a), it is inside an embedded declarative clause, while in (15-b), it is inside an embedded interrogative clause. In (15-c) and (15-d), *tinapay* is clefted. The contrast between (15-c) and (15-d) shows the following: *tinapay* can be clefted, as in (15-c), but movement out of the clause headed by "whether" is blocked, as in (15-d).

(15) a. Sinabi ko na binili ni Juan ang **tinapay**.
said.PV 1SG.GEN LNK bought.PV GEN PN NOM bread
'I said that Juan bought the bread.'

b. Sinabi ko kung binili ni Juan ang **tinapay**.
said.PV 1SG.GEN whether bought.PV GEN PN NOM bread
'I said whether Juan bought the bread.'

c. Ang **tinapay** ang sinabi ko na binili ni Juan.
NOM bread NOM said.PV 1SG.GEN LNK bought.PV GEN PN
'It is the bread that I said that Juan bought.'

d. *Ang **tinapay** ang sinabi ko kung binili ni Juan.
NOM bread NOM said.PV 1SG.GEN whether bought.PV GEN PN
Intended: It is the bread$_i$ that you wondered whether Juan bought it$_i$.'

Second, Tagalog is also sensitive to complex NP islands, as shown in (16). In (16-a) and (16-b), *guro* 'teacher' is in its canonical post-verbal position. The crucial difference between the two is that in (16-a), it is inside an embedded declarative clause, while in (16-b), it is inside a complex NP. In (16-c) and (16-d), *guro* is clefted. The contrast between (16-c) and (16-d) shows the following: *guro* can be clefted, as in (16-c), but movement out of the complex NP is blocked, as in (16-d).

(16) a. Pinaniwalaan ni Juan na minahal ni Pedro ang **guro**
believed.APPL GEN PN LNK loved.PV GEN PN NOM teacher
'Juan believed that Pedro loved the teacher.'

b. Pinaniwalaan ni Juan ang chika na minahal ni Pedro ang **guro**
believed.APPL GEN PN NOM gossip LNK loved.PV GEN PN NOM teacher
'Juan believed the gossip that Pedro loved the teacher.'

c. Ang **guro** ang pinaniwalaan ni Juan na minahal ni Pedro
NOM teacher NOM believed.APPL GEN PN LNK loved.PV GEN PN
'It is the teacher that Juan believed that Pedro loved.'

d. *Ang **guro** ang pinaniwalaan ni Juan ang chika na minahal ni
NOM teacher NOM believed.APPL GEN PN NOM gossip LNK loved.PV GEN
Pedro
PN
Intended: It is the teacher$_i$ about whom Juan believed the gossip that Pedro loved him/her$_i$

The examples in (15) and (16) suggest that Tagalog is an island-sensitive language. This claim is further corroborated by the ratings of the filler items associated with Experiments 2 and 3. See Table 3.13. Having established that Tagalog is island-sensitive independent of voice-mismatch extractions now puts us in a position to better evaluate whether a movement-based analysis of subject extractions under NAV is tenable. Below I use island effects to probe this question. Should these types of extractions exhibit insensitivity to island effects, then a movement-based analysis is less tenable. However, if they exhibit the same sensitivity as those in (15) and (16), then a movement-based analysis is more tenable.

Subject extractions under NAV are also island-sensitive

For speakers who allow subject extractions under NAV, these voice-mismatched extractions are also sensitive to *wh*-islands and complex NP islands in (17) and (18), respectively. For example, the subject can be extracted out of an embedded declarative clause that exhibits PV, as in (17-a). However, it cannot be extracted out of an embedded interrogative clause that exhibits PV, as in (17-b). Similarly, the contrast between (18-a) and

(18-b) suggest that the subject cannot be extracted out of a complex NP.

(17) a. Ang lalaki ang sinabi ko na binili ang **tinapay**.
NOM man NOM said.PV 1SG.GEN LNK bought.PV NOM bread
'It is the man that I said bought the bread.'

b. *Ang lalaki ang sinabi ko kung binili ang **tinapay**.
NOM man NOM said.PV 1SG.GEN whether bought.PV NOM bread
Intended: It is the man$_i$ that I said whether he$_i$ bought the bread.'

(18) a. Ang lalaki ang pinaniwalaan ni Juan na minahal ang **guro**
NOM man NOM believed.APPL GEN PN LNK loved.PV NOM teacher
'It is the man$_i$ that Juan believed that he$_i$ loved the teacher.'

b. *Ang lalaki ang pinaniwalaan ni Juan ang chika na minahal ang
NOM man NOM believed.APPL GEN PN NOM gossip LNK loved.PV NOM
guro
teacher
Intended: It is the man$_i$ about whom Juan believed the gossip that he$_i$ loved the teacher

If we take seriously this sensitivity to island effects as a diagnostic for movement, then subject extractions under NAV must either share the same derivations as their voice match counterparts. Or at the very least, they are similar enough in that they are also derived via A-bar movement.

A sketch of modifications for movement-based analyses

My goal here is to sketch out minimal changes to existing formal analyses so that we can allow voice mismatch extractions. I leave it for future work to flesh out the ramifications of these modifications.

Before proceeding, if we are committed to a movement-based analysis, then one obvious puzzle that we are confronted with is the lack of case-connectivity, as in (19). The agent *sundalo* 'soldier' is marked genitive in its canonical post-verbal position, as in (19-a).

However, when it is clefted, it surfaces with nominative marking, as in (19-b).

(19) Adapted from Ceña and Nolasco (2011, 248–9)

a. Pinasan ng sundalo ang **anak**
carried.PV GEN soldier NOM offspring
'The soldier carried his/her child.' *Baseline*

b. Ang sundalo ang pinasan ang **anak**
NOM soldier NOM carried.PV NOM child
'It's the soldier who carried his/her child.' *Cleft*

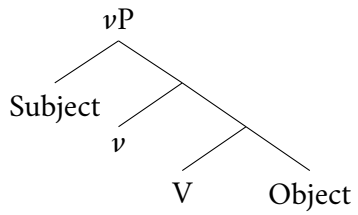
Following Lim and Erlewine (2018) who investigated multiple extraction in Bikol, an Austronesian language closely related to Tagalog, we can argue that morphological case can be structurally assigned *or* realized as context-sensitive defaults. Under this system of assigning morphological case, agreement with C/T assigns nominative case to its goal. In AV, T assigns nominative case to the subject, while in PV, it assigns patient case to the object. Under this system, other arguments are licensed depending on where they are in the structure. The default case in the *v*P is genitive, following Erlewine, Levin, and van Urk (2018). The default case outside the *v*P is nominative. This way of assigning morphological case could allow us to explain why an argument that is marked genitive at the tail of an A-bar chain can be spelled out nominative elsewhere.

Recall that I surveyed four recent approaches to voice morphology and how they account for voice match extractions in Tagalog in Chapter 2. Despite the differences in the finer details of these proposals, they can be (roughly) classified into two classes of analyses: (i) those involving multiple specifiers; and (ii) those that treat voice morphology as agreement with C.

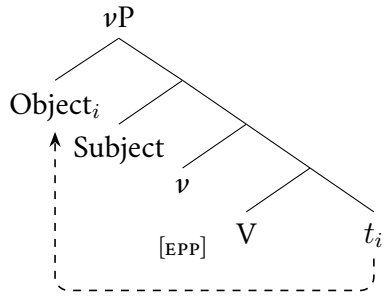
Approach 1: Multiple specifers. Recall that AV and non-AV voices are structurally different. In NAV (i.e., PV, LV, etc), the highest internal argument vacates the XP-sister of

*v*P—VP for PV and AppP for the other NAVs—and moves to the outermost specifier of *v*P, above the subject. The structures are provided in (20).

(20) a. Partial derivation of AV

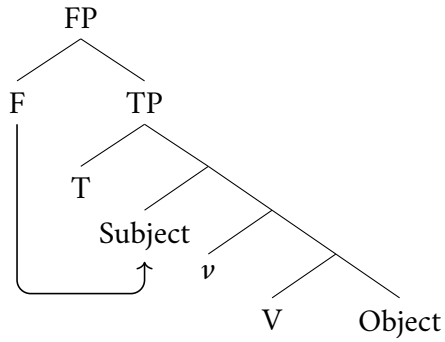


b. Partial derivation of PV



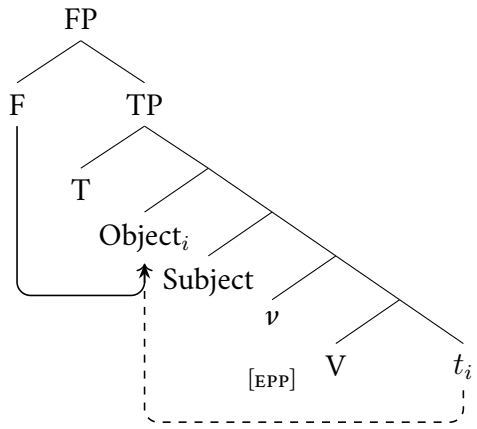
Also recall that voice-match extraction is captured by appealing to a phase-based theory of extraction, supplemented by some version of Attract Closest. For example, as schematized in (21), when the verb exhibits AV, only the subject is at the phase-edge. Thus, if a probe *F* were to search for a goal, under a phase-based theory of extraction, only the subject is visible to that probe because it is the most structurally prominent DP, by virtue of being the only one in the specifier of *v*P. On the other hand, the patient DP is opaque to this probe. This captures the desired patterns of (un)grammaticality.

(21) Partial derivation of AV



As schematized in (22), when the verb exhibits PV, the object evacuates the VP and moves to the outermost specifier of *v*P. Thus, if a probe *F* were to search for a goal, under a phase-based theory of extraction *plus* some version of Attract Closest, the patient DP is the more prominent DP and is thus the closest goal visible to that probe. On the other hand, the subject is no longer the most prominent DP at the edge of the *v*P-phase, rendering it opaque to this probe. This captures the desired patterns of (un)grammaticality patterns.

(22) Partial derivation of PV

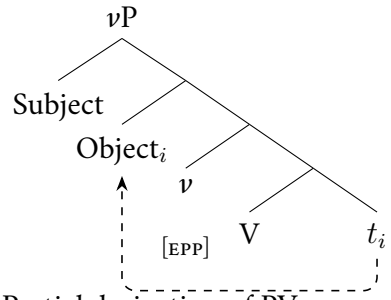


There are least two modifications that can allow us to account for subject extractions under NAV—assuming that for every option, everything else is kept equal.

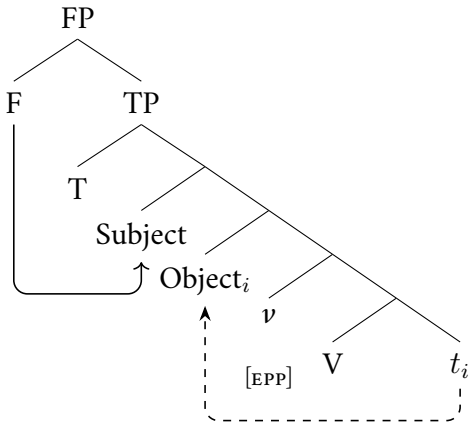
Implicit in these derivations is the "no tucking-in" assumption. Thus, one modification could be to relax this assumption. That is, in some NAV-derivations, some speakers

do tuck-in the object when it moves to the outermost specifier *v*P to satisfy [EPP] on *v*, as in (23-a). Thus, when a higher probe *F* searches a goal in its domain, it will be able to see the subject, as in (23-b).

(23) a. Partial derivation of PV



b. Partial derivation of PV



Object extractions under AV are still disallowed under this system. When the verb has AV, only the subject is in at the phase-edge. Thus, if a probe *F* were to search for a goal, the object is not visible.

If we say that the locus of variation resides in whether speakers tuck-in or not, an open question would be how to allow voice match extractions for those who do tuck-in. After all, a speaker that allows voice mismatch extractions still uncontroversially allows voice match extractions. If we say that those who allow voice mismatch extractions have access to two derivations for NAVs, one where the object is tucked in and one where it is not, then the question becomes what determines when a speaker will have a tucked-in

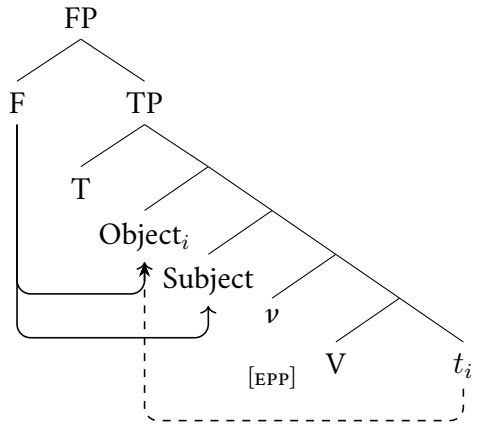
derivation and when she will have one that is not.

A different modification would be to invoke EQUIDISTANCE, whose definition is provided in (24), for some speakers.

(24) Equidistance: Elements in the same immediate maximal projection are equidistant from any other position.

In other words, the locus of variation could be that for some speakers, multiple specifier configurations are equidistant to a probe, and for others, distance is calculated using some version of Attract Closest. Let's consider how a speaker who has equidistant specifiers derive voice match and voice mismatch extractions. As in (25), when a higher probe *F* searches a goal in its domain, it will be able to see both the subject and the object. If *F* decides to pull up the object, then this will be an instance of voice match extraction. If, on the other hand, *F* decides to pull up the subject, then this will be an instance of voice mismatch extractions.

(25) Partial derivation of PV



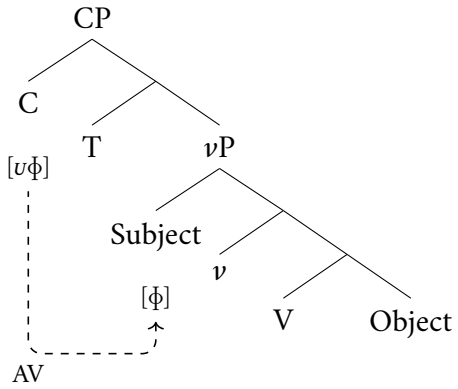
Object extractions under AV are still disallowed under this system. AV is a single-specifier structure. That is, only the subject is in its specifier. Thus, if a probe *F* were to search for a goal, the subject is the closest goal by virtue of being the only goal in the specifier of *v*P.

For present purposes, I simply asserted that equidistance can exist in our formal system. One open question for this modification is how we should bake in equidistance into our formal machinery. It is outside the scope of this dissertation to give a detailed treatment of how this can be done. This could be a stipulation à la Chomsky (2000, p. 122, 130). However, see the discussion by Hornstein (2009, pp. 43–4) on how equidistance in multiple specifier configurations is not a stipulation because they are only apparent exceptions to minimality. Or equidistance could be a property that emerges from several otherwise independently attested syntactic operations à la Longenbaugh and Polinsky (2018). I remain agnostic as to how this should be done.

Approaches #2: Agreement with C. Recall that under this umbrella of proposals, voice morphology is the spell-out when C enters into an Agree relation with whatever DP surfaces with *ang*.

Let us first consider the ϕ-agreement analysis (Aldridge, 2018). When C probes down its search space to value its unvalued feature, [*u*ϕ], it sees the subject with the feature [ϕ] that could value [*u*ϕ]. C Agrees with this DP and the result is AV, as schematized in (26).

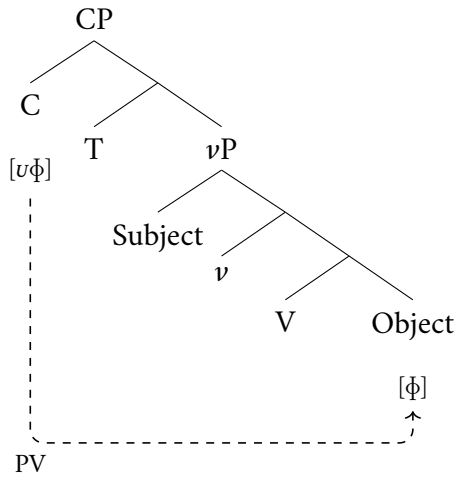
(26) Partial derivation of AV



When C probes down its search space to value its unvalued feature, [*u*ϕ], it sees the object DP with the feature [ϕ] that could value [*u*ϕ]. C Agrees with this DP and the result is PV,

as schematized in (27). Under Aldridge's system, C skips over the subject because *v* had already assigned inherent case to it via spec-head configuration.

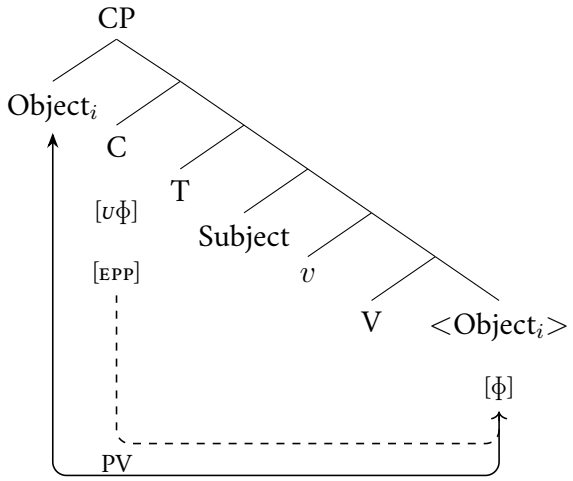
(27) Partial derivation of PV



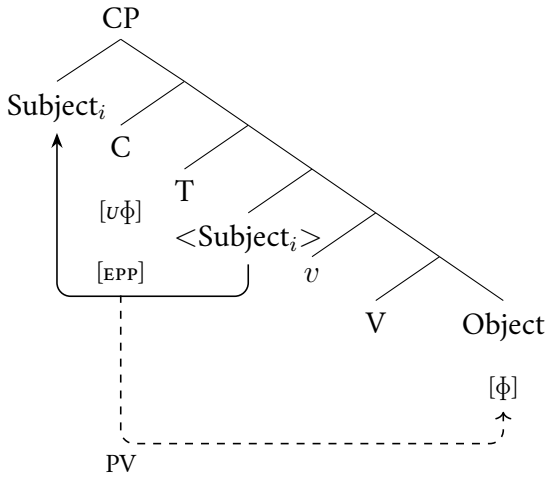
The voice match extraction restriction is enforced by singling out C, which can have an [EPP], as the only driver of movement in the language. If the DP also checks the [EPP]-feature during the Agree-relation shown in (26) for AV and (27) for PV, then the DP moves to the specifier of CP. Framed differently, movement is only possible for the DP with which C has entered into an Agree relation.

One modification that can allow us to account for subject extraction under NAV is to divorce the licensing of [*u*ϕ] and [EPP] on C. In other words, these licensing requirements need not be satisfied by the same DP. In PV, the object can satisfy [*u*ϕ] and then the subject can satisfy the [EPP] on C. The derivation in (28) shows canonical object extraction under PV, when the same DP licenses both [*u*ϕ] and the [EPP] on C. The derivation in (29) shows subject extraction under PV, when different DPs license [*u*ϕ] (checked off by the object) and the [EPP] (checked off by the subject) on C.

(28)



(29)



Aldridge (2018) has already baked in something akin to this in her proposal in order to account for adjunct extraction, as in (30), and for possessor extraction, as in (31-b): "the EPP and licensing properties of the [*u*ϕ] feature on C can be satisfied by separate constituents". In (30), [*u*ϕ] is licensed by the object *libro* 'book' while [EPP] is checked off by moving *saan* 'where.' In (31), [*u*ϕ] is licensed by the entire object *sungay ng kalabaw* 'horn of the water buffalo' while [EPP] is checked off by moving only a sub-part of what C Agreed with, *kalabaw*.

(30) Saan mo binili ang **libro**?
where 2SG.GEN boughtPV NOM book
'Where did you buy the book?' *Adjunct extraction*

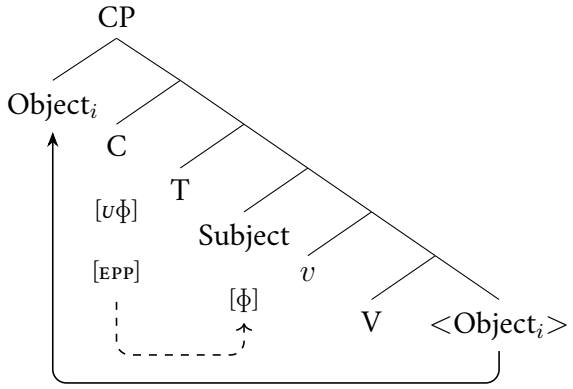
(31) Adapted from Kroeger (1993, p. 32)

a. Pinutol ng magsasaka ang **sungay ng kalabaw**
cut.PV GEN farmer NOM horn GEN water buffalo
'The farmer cut off the water buffalo's horn.' *Baseline*

b. **Ang kalabaw**, pinutol ng magsasaka ang **sungay**
NOM water buffalo cut.PV GEN farmer NOM horn
'The buffalo, the farmer cut off its horn.' *Possessor-extraction*

If we allow different constituents to satisfy the [$u\phi$] and [EPP] on C, then an open question is how to exclude object extractions under AV and under applicatives, which are uncontroversially bad. These extractions have the following in common: the argument that is extracted (i.e., that which checks off [EPP] on C) is lower in the structure than what determines voice (i.e., that which satisfies [$u\phi$]). Consider the scenario schematized in (32). [$u\phi$] could be licensed by the subject and [EPP] could be checked off by moving the object.

(32)



There are multiple ways to exclude this. One way is to stipulate that only the DP that is structurally higher than the DP that licenses [$u\phi$] can satisfy [EPP]. This effectively narrows down what could be used to satisfy the [EPP] to just the subjects. However, this stipulation is problematic because it makes the prediction that VP-level adjuncts are able to extract when the verb exhibits AV, but unable to when the verb exhibits PV or other NAVs. As seen in (33), this prediction is simply not borne out empirically. A VP-level adjunct like

kailan 'when' is able to extract, irrespective of voice.

(33) Adjunct extractions are possible under AV and NAV

a. Kailan nagbeyk ng keyk si **Maria** para kay Pilar?
when baked.AV GEN cake NOM PN for DAT PN
'When did Maria bake a cake for Pilar?'

b. Kailan bineyk ni Maria ang **keyk** para kay Pilar?
when baked.PV GEN PN NOM cake for DAT PN
'When did Maria bake the cake for Pilar?'

c. Kailan bineykan ni Maria si **Pilar** ng keyk?
when baked.APP GEN PN NOM PN GEN cake
'When did Maria bake a cake for Pilar?'

A second way is to play with the relative timing of probing. For example, C could probe for a goal to satisfy its [$u\phi$]. This probing cycle gives us PV. Then, C probes again to satisfy [EPP]. This probing cycle moves the first DP it encounters that is eligible to be moved. In NAVs, this would mean the subject. This also blocks object extraction when the verb has AV because during the second probing cycle, when C probes again to satisfy [EPP], C will always encounter the subject first. How to effectively exclude object extractions under AV and under applicatives remains an open question.

If we allow different constituents to satisfy the [$u\phi$] and [EPP] on C in our system, then we can frame the locus of variation in terms of where speakers draw the line of when C probes together (i.e., [$u\phi$] and [EPP] must be licensed by the same constituent) and when C probes separately (i.e., [$u\phi$] and [EPP] can be satisfied by different constituents). Table 3.17 summarizes when C probes together or separately in Tagalog per Aldridge (2018) and the current modification.

Now let's consider the bundling analysis (V. Chen, 2017). Recall that AV is the spell-out when C, which has a [$u\delta$], and T, which has [$u\phi$], both target the same goal, the subject,

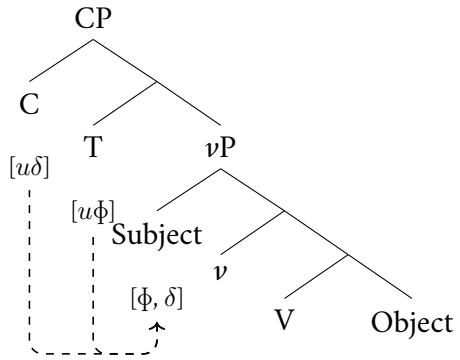
Table 3.17: A summary of when C probes together or separately in Tagalog under a ϕ-agreement approach to voice and the extraction restriction

	Voice match extractions	Voice mismatch extractions	Adjunct extraction	Sub-extraction
Spkr 1				
Spkr 2				

N.B. In grey are the dependencies where C probes separately

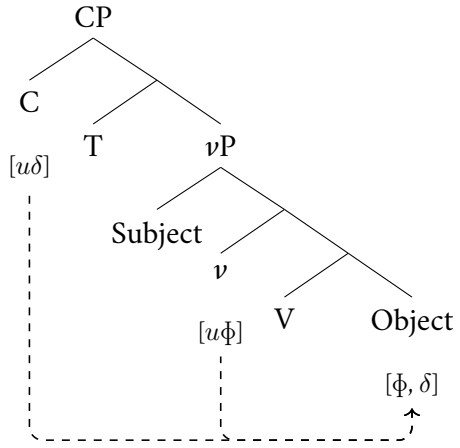
to value their respective features. This is schematized in (34).

(34) Partial derivation of AV



PV is the spell-out when C, which has a [*u*δ], and *v*, which has [*u*ϕ], both target the same goal, the object, to value their respective features. This is schematized in (35).

(35) Partial derivation of PV



Voice match extraction restriction is enforced by singling out C, which can have an [EPP], as the only driver of movement in the language. If the DP also checks the [EPP]-feature during the Agree-relation shown in (34) for AV and (35) for PV, then the DP moves to the specifier of CP. Framed differently, movement is only possible for the DP with which C and the other functional head responsible for voice (i.e., T for AV or *v* for PV) have entered into an Agree relation.

At first blush, it seems difficult to reconcile how the bundling analysis can account for voice mismatch extractions since under this view, determining voice morphology and enforcing voice match extraction restriction are a product of the same process: two probes Agreeing with the same goal. However, they need not be a single process. They could be separate but concurrent processes instead. This decoupling of determining the voice and enforcing the extraction restriction sounds much like divorcing the licensing of [$u\phi$] and [EPP] on C under the ϕ-agreement analysis, as discussed above. Thus, the modification for the bundling analysis will be in the same spirit as the modification discussed above for ϕ-agreement analysis.

3.7.3 Implications for sentence processing

Consider the received view of the extraction restriction discussed in Chapter 2: only the noun that is cross-referenced by voice is eligible to move (i.e., undergo syntactic operations like relativization, *wh*-question formation, topicalization, etc). Because of the one-to-one mapping between voice morphology and what is considered to be a licit interpretation of a FGD, it is reasonable to hypothesize that voice morphology allows Tagalog comprehenders to commit to an interpretation as early as the verb. Upon encountering the verb, comprehenders already have information about the event structure and by virtue of the voice morphology, the thematic role of the filler. Thus, at the verb, comprehenders should already have enough information to evaluate the thematic fit of the filler.

The hypothesis presented above crucially hinges on the assumption that comprehenders use this very coarse-grained information about the intimate one-to-one correlation between the identity of the gap and the particular form of voice morphology. In other words, the hypothesis assumes the following conditional probabilities of a gap corresponding to a particular form of voice morphology in (36).

(36) a. $P(\textsc{gap} = \text{Subject} | \textsc{voice} = \text{AV}) \sim 1$

b. $P(\textsc{gap} = \text{Object} | \textsc{voice} = \text{PV}) \sim 1$

As discussed in this chapter, there may be reasons to believe that the extraction restriction is not clean-cut. We saw in Experiment 1 that subject extractions under PV were given intermediate ratings, which is consistent with what Ceña and Nolasco (2011, 2012) documented. We saw in Experiments 2 and 3 that these findings are replicable and can be extended to the other voices in the language. If comprehenders attended to finer-grained information like the more nuanced conditional probabilities of a gap corresponding to a particular form of voice morphology in (37), which are based loosely on the results of Experiments 1–3, then this may have potential consequences for how voice could be used as cue in real-time. Because AV-morphology was a strong diagnostic of a subject gap, we might expect AV to be a robust trigger for dependency formation in *all* environments, as previously hypothesized. In contrast, because PV-morphology was less diagnostic of an object gap, we might expect PV to trigger less predictive parsing. More specifically, we might expect that comprehenders would be the most predictive in *wh*-questions, then least predictive in *ay*-inverted sentences and RCs. I summarize these potential consequences in (37).

(37) a. $P(\textsc{gap} = \text{Subject} | \textsc{voice} = \text{AV}) \sim 1$

b. $P(\textsc{gap} = \text{Object} | \textsc{voice} = \text{PV}, \textsc{fgd} = \text{WHQ}) \sim 1$

$$P(\textsc{gap} = \text{Object} | \textsc{voice} = \text{PV}, \textsc{fgd} = \text{AY}) < 1$$

$$P(\textsc{gap} = \text{Object} | \textsc{voice} = \text{PV}, \textsc{fgd} = \text{RC}) < 1$$

Chapter 4

The role of voice in dependency formation

The main goal of this chapter is to better understand the role of Tagalog voice morphology in the incremental processing of FGDs. I asked the following questions:

1. Is voice morphology a cue that Tagalog comprehenders use in real-time?
2. If it is, are the different types of voice used the same way?
3. Is voice used the same way across different types of dependencies?

These questions can be classified into two categories. The first is the WHETHER-question, and the last two are the HOW-questions. Before we can ask the how-questions, it is imperative that we first ask the whether-question. As I had mentioned in Chapter 1, the whether-question is always an empirical question. What we know from the grammatical illusion literature is that comprehenders are sensitive to cues that are not even instantiated by the dependency of interest. What we know from the acquisition literature is that comprehenders can also ignore very reliable cues and at times, even do so in favor of less reliable ones. Thus, even if voice morphology is a cue that is available to Tagalog

comprehenders in the input and as discussed in Chapters 2 and 3, is generally reliable in determining the identity of the filler in elicitation sessions and untimed judgment tasks, whether they actually use it in real-time remains an open question. How they will use it is another open question.

In this chapter, I establish that voice morphology is a cue that Tagalog comprehenders attend to and use when processing FGDs in real-time. I add to the existing literature supporting the view that comprehenders actively construct dependencies even in the absence of disambiguating information. Crucially, I do so using a language that is heavily under-represented in the psycholinguistics literature. I also provide evidence that the way in which comprehenders use voice as a cue may be modulated by different classes of linguistic information, including (but not limited to) finer-grained information about the conditional probabilities of a gap corresponding to a particular form of voice morphology, and information about the structural similarities or the thematic complexity of the constructions involved and the relative frequency of the different types of voice. The interactions of these classes of information lead to processing asymmetries.

This chapter is organized as follows. Section 4.1 reviews the literature on active dependency formation, that is, the idea that comprehenders do not wait until they encounter all of the bottom-up information to start interpreting a sentence. Section 4.2 then presents a hypothesis about how Tagalog comprehenders could use information encoded by voice morphology to guide their interpretations when processing FGDs in real-time. Section 4.3 is an attempt to answer the whether-question posted above. Using a stops-making-sense task (Boland et al., 1995), I compared the time course of dependency formation when the verb exhibited voice and when it did not to approximate the independent contribution of voice in FGD-formation. Section 4.4 is an attempt to answer the how-questions posted above. Using the same methodology, I compared the time course of dependency formation when the verb exhibited AV and when it exhibited PV across three types of

FGDs—*wh*-questions, RCs, and *ay*-inverted sentences—syntactic constructions surveyed back in Chapter 2. Finally, in Section 4.5, I conclude by discussing the level of granularity of the voice cue comprehenders use in real-time, and by examining the various asymmetries in their behavioral data to see how classes of information can interact.

4.1 Comprehenders actively construct dependencies

FGDs challenge comprehenders to associate the filler with a gap, whose position is not always unambiguously indicated by the evidence in the input. Nevertheless, comprehenders do not wait until all of the bottom-up information becomes available (Crain & Fodor, 1985; L. Frazier, 1987; L. Frazier & Clifton, 1987). Instead, they attempt to establish this link by predictively positing a gap at each available position that would allow the dependency to be resolved without violating more abstract grammatical constraints like islandhood (Traxler & Pickering, 1996; Wagers & Phillips, 2009).

One piece of evidence supporting active dependency formation is the FILLED-GAP EFFECT (Crain & Fodor, 1985; Stowe, 1986). For example, Stowe (1986) found that reading times at the direct object pronoun *us* were longer in embedded *wh*-questions, like (1-a), than in minimally different embedded polar questions, like (1-b). This suggests that comprehenders posited a gap in direct object position after encountering *bring* in (1-a), without waiting to check whether it was already occupied. Upon encountering *us*, their initial predictions were disconfirmed, as indicated by the longer reading times at this region.

(1) a. My brother wanted to know who Ruth will bring us home to ___ at Christmas.
b. My brother wanted to know if Ruth will bring us home to Mom at Christmas.

Another piece of evidence for active dependency formation is the PLAUSIBILITY MISMATCH

EFFECT (Boland et al., 1995; Garnsey, Tanenhaus, & Chapman, 1989; Traxler & Pickering, 1996). For example, Boland et al. (1995) found that participants judged the filler *which prize* in (2) to be implausible upon encountering the verb *visit* using the SMS paradigm, a self-paced word-by-word reading combined with an explicit acceptability judgment task. The filler *which prize* is implausible only if it is associated with the verb *visit*, as in (2-a). However, it can be plausible if it is associated not with *visit* but with a different verb downstream, as in (2-b). The comprehenders' judgment suggests that they were attempting to resolve the dependency prior to the identification of the gap's location.

(2) a. #Which prize did the salesman visit ___ while in the city?

b. Which prize did the salesman visit Chicago to obtain ___?

There is a wealth of cross-methodological evidence that comprehenders actively construct FGDs: electrophysiology using EEG (Garnsey et al., 1989), the eye-tracking while reading paradigm (Traxler & Pickering, 1996), eye-movements in the visual world paradigm (Sussman & Sedivy, 2003), self-paced reading (Lee, 2004), among many others. There is also evidence that active dependency formation is cross-linguistically robust: Bangla (Chacón et al., 2016), Chamorro (Wagers et al., 2015), Dutch (L. Frazier & Flores d'Arcais, 1989), German (Schlesewsky, Fanselow, Kliegl, & Krems, 2000), Italian (Vincenzi, 1991), Japanese (Aoshima, Phillips, & Weinberg, 2004), and Russian (Sekerina, 2003), to name a few. There is also evidence that it is robust in both neurotypical and aphasic populations (Dickey, Choy, & Thompson, 2007).

Actively associating a filler to a gap may be risky because the initial association may be incorrect and thus, needs revision. How can comprehenders minimize these faulty associations? Framed differently, what types of linguistic information do comprehenders use to guide their incremental structure-building? It is perhaps unsurprising that comprehenders can use different classes of linguistic information. They can leverage syntactic

information, like word order (Stowe, 1986) or more abstract structural constraints like islandhood (Traxler & Pickering, 1996; Wagers & Phillips, 2009). They can use lexical information, like a verb's argument structure and thematic relations (Trueswell, Tanenhaus, & Garnsey, 1994). They can also use morphological information, like case (Kamide, Altmann, & Haywood, 2003), and prosodic information (Nagel, Shapiro, & Nawy, 1994).

More directly related to the questions that this chapter is asking is the investigation by Wagers et al. (2015) of what is called *wh*-agreement in the generative syntax literature and its contribution in processing FGDs in Chamorro, an Austronesian language distantly related to Tagalog. When present, *wh*-agreement provides direct morphological evidence for the presence of a FGD and for the grammatical function of the gap (Chung, 1982, 1994, 1998). Provided in (3) are examples involving the verb *låksi* 'to sew.' The infix *-in-* in (3-a) signals to Chamorro comprehenders that the filler *håfa na magågu* 'which clothes' is the object of the verb *låksi*. Meanwhile, when the verb only exhibits subject-verb agreement (which is different from *wh*-agreement), the filler *håfa na magågu* can either be the object of the verb, as in (3-b), or a possessor of the object of the verb, as in (3-c). Only after encountering *manggas-ñiha* downstream can these possibilities be disambiguated.

(3) Adapted from Wagers et al. (2015, p. 113)

a. Håfa na magågu linaksem-mu gi paingi
what LNK clothes WH[OBJ].SEW-AGR LOC last night
'Which clothes did you sew last night?' *With wh-agreement*

b. Håfa na magågu un låksi gi paingi
what LNK clothes AGR sew LOC last night
'Which clothes did you sew last night?' *Without wh-agreement*

c. Håfa na magågu un låksi gi paingi manggas-ñiha
what LNK clothes AGR sew LOC last night sleeves-AGR
'Which clothes did you sew sleeves of last night?' *Without wh-agreement*

Wagers et al. found that this piece of morphological information modulated how ac-

tively comprehenders constructed dependencies. Only when the verb exhibited overt *wh*-agreement did the comprehenders actively construct FGDs. When the verb could have *wh*-agreement but did not, dependency formation was delayed.

Here I ask how Tagalog comprehenders use similar morphological information to construct FGDs. See Chapter 2 for the empirical generalizations. Like Chamorro, Tagalog exhibits rich verbal morphology (i.e., voice morphology) that is intimately correlated with which argument can participate in such dependencies. Unlike Chamorro, this piece of morphology appears in both FGD- and non-FGD contexts.

Since post-verbal word order is relatively free, it is difficult to ascertain where the exact location of the gap is. What is crucial is that voice morphology always precedes the gap—at least in the dependencies examined in this chapter. This early availability of voice, along with the other classes of information encoded on the verb, could potentially be leveraged by comprehenders in real-time FGD-processing. In what follows, I present a generalized schema of how voice could be used by an incremental comprehender and provide a walkthrough of how this schema applies by looking at examples of plausible and implausible extractions in the language.

4.2 A hypothesis: Voice allows comprehenders to commit to an interpretation upon encountering the verb

First consider the received view of the extraction restriction discussed in Chapter 2: only the noun that is cross-referenced by voice is eligible to move (i.e., undergo syntactic operations like relativization, *wh*-question formation, topicalization, etc). Because there is a one-to-one mapping between voice morphology and what is considered to be a licit interpretation of a filler-gap dependency, I hypothesize that voice morphology allows Tagalog comprehenders to commit to an interpretation as early as the verb. Upon en-

countering the verb, comprehenders already have information about the event structure and by virtue of the voice morphology, the grammatical/thematic role of the filler. At this point, they should have enough information to evaluate the thematic fit of the filler.

In Figure 4.1, I provide a generalized schema of how an incremental comprehender could use voice to commit to an interpretation upon encountering the verb in a FGD-context. In panel A, an incremental comprehender recognizes that a FGD is involved. There are various cues that she could use to infer this: (i) word order: she had just encountered a noun phrase before a verb in a verb-initial language; and (ii) overt markers of FGDs: immediately after the noun phrase she had just encountered, she may encounter a particle—like another *ang*, for example—that implicates a FGD. In this case, the second *ang* signals an upcoming cleft (i.e., *It was X that...*). In panel B, she encounters the verb with the voice morphology. In panel C, she projects a gap-site based on the information provided by voice along with the sub-categorization information provided by the verb itself. Since post-verbal word order is relatively free, it is difficult to ascertain where the exact location of the gap is. Her projection therefore need not be about its exact location but rather about the features that the gap must have to ensure a successful parse. Finally, in panel D, she encounters all the bottom-up information in the form of the co-argument that confirms or disconfirms her expectation.

Next, I look at examples of plausible and implausible extractions to see how the schema works. Approximate time points in a parse are indicated by circled numerals, which correspond to when the comprehender encounters crucial cues. Let's consider (4).

(4) Tinanong ni Maria kung ① ang bata ang ② bumili kahapon ③ ng laruan.
asked.PV GEN PN whether NOM child NOM bought.AV yesterday GEN toy
'Maria asked whether it was the child that bought a toy yesterday.'

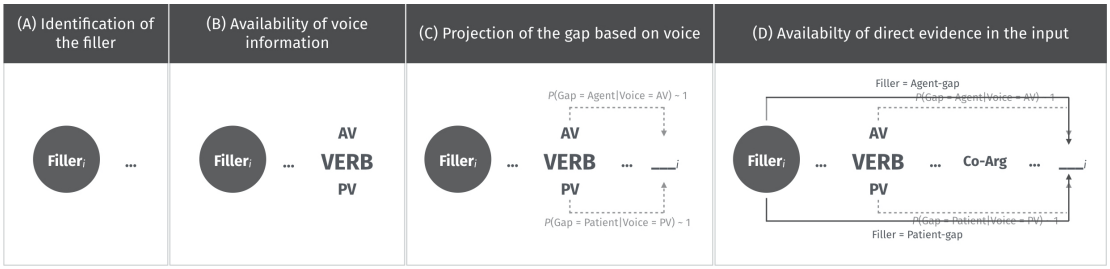


Figure 4.1: A schema of how a comprehender could use voice to guide her emerging interpretation. In (A), an incremental comprehender recognizes that she is in an FGD. In (B), she encounters the verb with the voice morphology. In (C), she projects a gap-site based on the information provided by voice along with the sub-categorization information provided by the verb itself. In (D), she encounters all the bottom-up information in the form of the co-argument that confirms or disconfirms her prediction

At the complementizer *kung* (before time point ①), the comprehender knows that she is in an environment that hosts embedded questions. Whether there will be a FGD involved remains unclear since *kung* is compatible with both polar (yes-no) and *wh*-questions. At time point ①, upon encountering the second *ang*, she knows that she is in a cleft construction and thus, there is a gap downstream that must be licensed. Moreover, she has clear evidence about what the filler is, *ang bata* 'the child.' At this point, the gap can be linked to the filler, although she has no evidence about its role in the sentence yet. She knows that there are unresolved requirements: (i) to associate the gap to a position with a specific thematic role, and (ii) if applicable, to have a co-argument to satisfy the verb's selectional requirements.

At time point ②, encountering the verb instantiates the event structure and crucially, by virtue of the voice morphology, the thematic role of the filler. She sees that the verb is semantically transitive and thus, needs to have another argument downstream. She also sees that the verb exhibits AV and consequently, projects a gap corresponding to the subject/agent of the verb, and resolves the open requirement from time point ① by linking the gap to the agent of *bili* 'buy'. Alternatively, she can be conservative and wait until the co-argument is encountered before making any commitments in order to reduce the risk

of faulty associations. At time point ③, upon encountering *laruan* 'toy,' she satisfies the subcategorization requirement of the verb. If she was conservative at time point ②, she now has all the information about how to link the filler and the gap.

Provided in (5) is a sentence that is minimally different from (4). This is a semantically anomalous sentence, as indicated by the pound sign ('#') in front. Instead of encountering the filler *ang bata* 'the child' at time point ①, she encounters *ang laruan* 'the toy.'

(5) #Tinanong ni Maria kung ① ang lauran ang ② b**um**ili ...
asked.PV GEN PN whether NOM toy NOM bought.AV ...
'Maria asked whether it was the toy that bought ... '

At time point ②, she sees that the verb is semantically transitive and thus, needs to have another argument downstream. She also sees that the verb exhibits AV and consequently, projects a gap corresponding to the agent, and resolves the open requirement from time point ① by linking the gap to the agent of the verb. Alternatively, she can be conservative and wait. If she does link the filler with the gap, then this should lead to a plausibility mismatch effect. The filler 'toy' is a poor fit for a gap corresponding to the agent of 'buy.'

The hypothesis presented above hinges on the assumption that comprehenders use this very coarse-grained information about the intimate correlation between the identity of the gap and the particular form of voice morphology. In other words, I am assuming that there is a one-to-one mapping between voice morphology and what is considered to be a licit interpretation of FGDs. I am also assuming that voice is used consistently across different types of dependencies. My assumptions are summarized in (6).

(6) a. $P(\textsc{gap} = \text{Subject/Agent} | \textsc{voice} = \text{AV}) \sim 1$

b. $P(\textsc{gap} = \text{Object/Patient} | \textsc{voice} = \text{PV}) \sim 1$

As discussed in Chapter 3, there are reasons to believe that the voice-match extraction

restriction might not be as robust as previously documented. We saw in Experiment 1 that subject extractions under PV were given intermediate ratings, which is consistent with what Ceña and Nolasco (2011, 2012) documented. We saw in Experiments 2 and 3 that these findings are replicable and can be extended to the other voices in the language. That the extraction restriction is not as clean-cut as was originally believed may have potential consequences for how voice could be used as cue in real-time. Because AV-morphology was a strong diagnostic of a subject-gap, we might expect AV to be a robust trigger for dependency formation in *all* environments, as previously hypothesized. In contrast, because PV-morphology was less diagnostic of an object-gap, we might expect PV to trigger less predictive parsing. We might expect that comprehenders would be the most predictive in *wh*-questions, then least predictive in *ay*-inverted sentences and RCs. I summarize these potential consequences in (7).

(7) a. $P(\textsc{gap} = \text{Subject/Agent} | \textsc{voice} = \text{AV}) \sim 1$

b. $P(\textsc{gap} = \text{Object/Patient} | \textsc{voice} = \text{PV}, \textsc{fgd} = \text{WHQ}) \sim 1$

$P(\textsc{gap} = \text{Object/Patient} | \textsc{voice} = \text{PV}, \textsc{fgd} = \text{AY}) < 1$

$P(\textsc{gap} = \text{Object/Patient} | \textsc{voice} = \text{PV}, \textsc{fgd} = \text{RC}) < 1$

Before we can ask how comprehenders will leverage voice, we first need to establish whether they attend to this cue and actually use it. Cue availability and reliability do not necessarily entail use. In experiment 4, I address the whether-question by comparing the time course of dependency formation when the verb exhibited voice and when it did not. In experiment 5, I address the how-question by directly comparing the time course of dependency formation when the verb exhibited AV and PV.

4.3 Experiment 4: Isolating the contribution of voice

Experiment 4 was designed to isolate the independent contribution of voice when processing filler-gap dependencies. In other words, in Experiment 4, I ask the whether-question by comparing the time courses of dependency formation when the verb exhibited voice and when it did not. Tagalog allows for this comparison to be made because there are grammatical aspects where verbs do not have voice. Crucially, these voiceless paradigms impose comparable restrictions on displacement as when verbs have voice. These voiceless paradigms are discussed in section 4.3.1.

I used a phrase-by-phrase non-cumulative moving window SMS paradigm (Boland et al., 1990, 1995) in order to detect plausibility mismatch effects that arise when comprehenders link fillers and gaps in *d*-linked *wh*-questions, which are *wh*-questions that have the shape *which N* in English. In this paradigm, sentences were presented one phrase at a time. The phrases did not accumulate across the screen. The participant controlled the presentation rate by pressing the F-key after reading each phrase. The keypress caused the next phrase to appear, unless she pressed the J-key to indicate that the sentence stopped making sense. The participant sets her own criterion for when she will reject the sentence, and following Boland et al. (1995), my linking hypothesis is that plausibility mismatches increase the likelihood that she will press the J-key at a given phrase. Implausible sentences will be nonsensical in a way that only becomes apparent when the participants are able to integrate all of the linguistic material into a coherent whole. When they have indicated that a sentence stopped making sense, it can be inferred that they have assembled the pieces of that sentence into a meaningful unit at that particular point in time.

Like standard self-paced reading paradigm (SPR; Just, Carpenter, & Woolley, 1982), SMS forces the comprehender into a strictly incremental mode of processing. However, SMS has some advantages over SPR (Boland et al., 1995; Mauner, Tanenhaus, & Carlson, 1995; Mauner & Koenig, 2000): (i) SMS data has less variability across conditions com-

pared to SPR data; (ii) it is highly sensitive to anomalies involving the verb's argument structure; and (iii) plausibility effects emerge more locally, typically at the region where incongruity is introduced. Here I take advantage of this more localized estimate of where (i.e., at which region) comprehenders commit to an intended parse.

4.3.1 Voiceless verbs impose similar restrictions on movement

Even though verbs in Tagalog typically exhibit voice, there are aspectual paradigms in the language where they do not. Crucially, even though verbs in these grammatical aspects lack voice, the grammar of the language imposes comparable restrictions on extraction as those that do exhibit voice. Thus, comparing their time courses of dependency formation allows us to approximate the independent contribution of voice in FGD-processing.

The iterative aspect (glossed as ITER) is used to denote repeated or prolonged actions. Voice is optional in iteratives (Schachter & Otanes, 1983, pp. 398–9). What is crucial here is that only the extraction of the subject is licit when the verb does not exhibit voice, much like when the verb exhibits AV. Because AV and voiceless iteratives impose comparable restrictions on extraction, I argue that this pairwise comparison allows us to isolate the effect of AV in processing FGDs. Their comparability is exemplified in (8).

(8) a. Medyo mataba ang **bata** na b**um**ati ng mga bisita
somewhat fat NOM child LNK greet.AV GEN PL visitor
'The child who greeted the visitors was chubby.'

Intended: *The child who the visitors greeted was chubby.

b. Medyo mataba ang **bata** na bati nang bati ng mga bisita
somewhat fat NOM child LNK greet.ITER GEN PL visitor
'The child who kept on greeting the visitors was chubby.'

Intended: *The child who the visitors kept on greeting was chubby.

The recent perfective aspect (glossed as RP) is used to denote events that have occurred in

the recent past. This aspect exhibits two quirky properties: (i) the verb exhibits no voice; and (ii) both arguments typically receive the same case marking (Kroeger, 1993; M. Frazier & Yoshida, 2013). Schematically, a clause in the recent perfective aspect typically has the following shape: Verb – *ng* NP1 – *ng* NP2. While these properties conspire to make a sentence globally ambiguous, there is a robust preference for interpreting NP1 as the subject/agent, as in (9-a). Sauppe (2016) provided experimental evidence for this preference. Crucial here is the grammaticization of this preference when what immediately follows is a proper name or a pronoun, as in (9-b). The post-verbal argument can only be interpreted as the subject. Framed differently, in FGDs, the filler can now only be interpreted as the object.

(9) a. Medyo mataba ang **bata** na kakabati lang ng mga bisita
somewhat fat NOM child LNK greet.RP just GEN PL visitor
Preferred: 'The child who the visitors just greeted was chubby.'
OK: 'The child who just greeted the visitors was chubby.'

b. Medyo mataba ang **bata** na kakabati lang ni Pedro/niya
somewhat fat NOM child LNK greet.RP just GEN PN/3.SG.GEN
'The child who Pedro/he just greeted was chubby.'
Intended: *The child who just greeted Pedro/him was chubby.

I leveraged (the grammaticization of) this preference to isolate the contribution of PV in processing FGDs. As exemplified in (10), PV and recent perfectives now impose comparable restrictions on extraction, with the proviso that the co-argument in the recent perfective be a proper name or a pronoun.

(10) a. Medyo mataba ang **bata** na binati ni Pedro
somewhat fat NOM child LNK greet.PV GEN PN
'The child who the visitors greeted was chubby.'
Intended: *The child who greeted the visitors was chubby.'

b. Medyo mataba ang **bata** na kakabati lang ni Pedro
somewhat fat NOM child LNK greet.RP just GEN PN
'The child who Pedro just greeted was chubby.'

Intended: *'The child who just greeted Pedro was chubby.'

4.3.2 Participants

I recruited 80 speakers from UP – Diliman, using the same recruitment criteria discussed in Experiments 1–3. They ranged from 18 to 35 years of age (M = 23, SD = 4.5), and all lived in and around Metro Manila at the time of testing. They received a Starbucks gift certificate valued at 400 Philippine Pesos for participating.

4.3.3 Materials

The experiment employed a 2×2 design, crossing whether the verb exhibited voice or not (VOICE: +VOICE, −VOICE) and whether the filler was plausible or not (PLAUSIBILITY: +PLAUS, −PLAUS). I created 2 sets of 12 items involving non-reversible predicates—one for comparing the time course when the verb imposes an agent-only restriction on extraction (Experiment 4A), and one for comparing the time course when the verb imposes a patient-only restriction on extraction (experiment 4B). The same 12 verbs were used throughout the experiment but they appeared with different co-argument pairs. Each item was distributed evenly across four lists via Latin Square design. The experimental items were combined together in one session and randomized with 48 filler items that were invariant across all participants. The filler items involved other dependencies, half with voice morphology and half without, and were designed to deter participants from developing a strategy for when to provide a response. Half of the distractors had the plausibility mismatch effect arise before encountering the main verb, and the other half, after encountering the main verb. Provided in Table 4.1 are sample experimental items in the two sub-experiments.

Table 4.1: Sample item used in Experiment 4, manipulating VOICE (+, –) and PLAUS (+, –)

VOICE	PLAUS	Region					
		1	2	3	4	5	6+
		Aling	*dalaga*	*ang*	*umiinom parati*	*ng tubig*	...
+VOICE	+PLAUS	which	maiden		drink.AV always	water	
		Which young woman always drinks water ...?					
		Aling	*tubig*	*ang*	*umiinom parati*	*ng dalaga*	...
+VOICE	−PLAUS	which	water		drink.AV always	maiden	
		#Which water always drinks a young woman ...?					
		Aling	*dalaga*	*ang*	*inom nang inom*	*ng tubig*	...
−VOICE	+PLAUS	which	maiden		drink.ITER	water	
		Which young woman keeps on drinking water ...?					
		Aling	*tubig*	*ang*	*inom nang inom*	*ng dalaga*	...
−VOICE	−PLAUS	which	water		drink.ITER	maiden	...
		#Which young woman always drinks water ...?					

VOICE	PLAUS	Region					
		1	2	3	4	5	6+
		Aling	*alak*	*ang*	*ininom*	*niya*	...
+VOICE	+PLAUS	which	wine		drink.PV	3SG	
		Which wine did (s)he drink ...?					
		Aling	*babae*	*ang*	*ininom*	*niya*	...
+VOICE	−PLAUS	which	woman		drink.PV	3SG	
		#Which woman did (s)he drink (earlier) ...?					
		Aling	*alak*	*ang*	*kakainom lang*	*niya*	...
−VOICE	+PLAUS	which	wine		drink.RP just	3SG	
		Which wine did (s)he just drink ...?					
		Aling	*babae*	*ang*	*ininom*	*niya*	...
−VOICE	−PLAUS	which	woman		drink.RP just	3SG	
		#Which woman did (s)he just drink ...?					

4.3.4 Procedure

The experiment was presented using Linger (Rohde, 2003). Participants were first introduced to a moving window SPR to familiarize themselves with the presentation. They were instructed to press the F-key to see the first word of the sentence, and to press the F-key again to see the next word. They were to continue pressing the F-key until they saw the end of the sentence. After some practice SPR-presentations, they were introduced to the stops-making-sense task. They were instructed to read the sentences in the same way as before with one crucial difference: after each phrase, they had the option to abort the presentation by pressing the J-key if the sentence stopped making sense. Aborting the presentation led to the presentation of another sentence. After completing 4 practice items, they started the experiment proper.

4.3.5 Analysis

The analysis is couched in Signal Detection Theory (SDT; Macmillan & Creelman, 2005), a common framework for modeling decision-making processes under uncertainty. For the purposes of this chapter, the framework can be applied to a comprehender's judgments of whether to reject or continue with an item that may be plausible or implausible. A standard SDT analytical approach to the stops-making-sense task categorizes response behaviors into correct rejections of implausible items (i.e., hits) and incorrect rejections of plausible items (i.e., false alarms). The distributions of the hits and false alarms in a detection task can be used to construct an explicit model of the comprehender's decision process. The comprehender's responses are attributed to a combination of how clearly the comprehender is able to detect signal from noise (i.e., their sensitivity, which is represented by *d'*) and the threshold that they used to make these judgments (i.e., their bias). Thus, one advantage of couching my analysis in SDT is the comprehender's sensitivity is decoupled from their bias.

To compare the differences in the time course of interpreting sentences with voice morphology and sentences without, I used the sensitivity index in SDT, expressed in *d'*, by scaling the participants' correct rejections of implausible items against their incorrect rejections of plausible items.

At each region, I counted the number of trials where participants had indicated that the sentence stopped making sense. I then added the rejection counts at a particular rejection and the rejection counts at preceding regions. I used these cumulated rejection counts to calculate their cumulative *d'*. I then re-sampled with replacement at the participant-level to create a new sample of participants of the same size (n = 80), from which I derived their cumulative *d'* at each region. I repeated this procedure 100,000 times. For each replication, I calculated the mean cumulative *d'* of each condition at each region. I then used these bootstrapped values to estimate the 95% confidence intervals at each region (Macmillan, Rotello, & Miller, 2004).

At any given region, a *d'* of 0 implies that participants on average did not discriminate implausible sentences from plausible ones. In other words, the rate at which they correctly rejected implausible sentences was comparable to the rate at which they incorrectly rejected plausible ones. A positive *d'* implies that on average, they discriminated implausible sentences from plausible ones, with more correct rejections for the former than misses for the latter. Finally, a negative *d'* implies that on average, they discriminated implausible sentences from plausible ones, but in the opposite direction: the rate at which they incorrectly rejected plausible sentences was greater than the rate at which they correctly rejected implausible ones. In lieu of reporting p-values, we provide the 95% CI for each region. If the CI does not contain 0, then the *d'* at that region is significantly different from 0 at the alpha level of .05.

To determine whether the presence of voice significantly affected the participants' sensitivity, I subtracted the *d'* of the conditions without voice from the *d'* of the conditions

with voice at each region for each replication. I refer to this difference as $\Delta d'$. I then calculated the mean $\Delta d'$ and the 95% CI at each region. A positive $\Delta d'$ implies that on average, participants had greater sensitivity when the verb exhibited voice than when it did not. A negative $\Delta d'$ implies the reverse. A $\Delta d'$ of 0 implies that the presence of voice did not affect their sensitivity. The 95% CIs are interpreted the same way.

4.3.6 Results

The main findings of Experiment 4 are the following:

- Upon encountering the verb, Tagalog comprehenders actively associated the filler with the gap even before the disambiguating information that came in the form of the co-argument;
- Voice enhanced this linkage by allowing them to commit to a correct interpretation upon encountering the verb.

Reported in Figure 4.2 the participants' phrase-by-phrase mean cumulative *d'*s and the 95% CI by VOICE at each region.

Experiment 4A: Comparing AV and iteratives

There were no plausibility effects observed in the first two regions. At the *ang*-region, participants were rejecting plausible sentences more than implausible sentences: d' = -.63, 95% CI [-.98, -.19]. They showed comparable sensitivity when the verb exhibited voice and when it did not: $\Delta d'$ = .06, 95% CI [-.23, .38].

At the verb-region, participants were rejecting implausible sentences more than plausible ones: $\Delta d'$ = .94, 95% CI [.48, 1.46]. However, they showed even greater sensitivity when the verb exhibited voice compared to when it did not: $\Delta d'$ = .43, 95% CI [.07, .83].

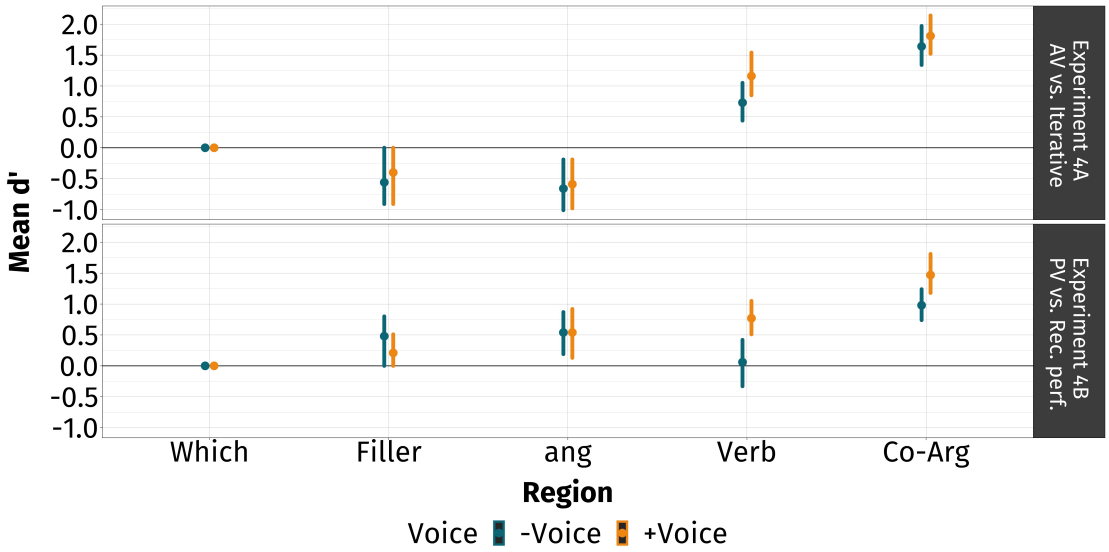


Figure 4.2: Phrase-by-phrase discrimination scores in experiment 4. Plotted are the mean scores at each region (points) with 95% confidence intervals (bands) derived by a bootstrap over participants. In blue are –VOICE-conditions and in gold are +VOICE-conditions. The top panel represents the comparison between AV and iteratives (experiment 4A), and the bottom panel represents the comparison between PV and recent perfectives (experiment 4B). The verb- and co-argument-regions are the critical and disambiguating regions, respectively.

At the co-argument region, participants were significantly rejecting implausible sentences more than plausible ones: d' = 1.73, 95% CI [1.39, 2.09]. They showed comparable sensitivity when the verb exhibited voice and when it did not: $\Delta d'$ = .17, 95% CI [-.14, .48].

Experiment 4B: Comparing PV and recent perfectives

There were no plausibility effects observed in the first two regions. At the *ang*-region, participants were rejecting implausible sentences more than plausible sentences: $\Delta d'$ = .54, 95% CI [.18, .89]. They showed comparable sensitivity when the verb exhibited voice and when it did not: $\Delta d'$ = .00, 95% CI [-.26, .26].

At the verb-region, participants were rejecting implausible sentences more than plausible ones when the verb exhibited voice but not when it did not: d' = .77, 95% CI [.51, 1.05]

for PV and d' = .06, 95% CI [-.33, .42] for recent perfectives. This is further corroborated by a significant difference in their d's: $\Delta d'$ = .71, 95% CI [.30, 1.16].

At the co-argument-region, participants were rejecting implausible sentences more than plausible ones: d' = 1.23, 95% CI [.78, 1.75]. However, they showed greater sensitivity when the verb exhibited voice than when it did not: $\Delta d'$ = .49, 95% CI [.15, .86].

4.3.7 Discussion

The goal of experiment 4 was to isolate the independent contribution of voice morphology in processing filler-gap dependencies by comparing the time courses of dependency formation when the verb exhibited voice and when it did not. I found the following. First, upon encountering the verb, participants were actively associating the filler with the gap. This association happened even before they had encountered fully disambiguating information in the form of the co-argument. Second, voice was used as an additional cue to further strengthen the participants' commitment to a correct interpretation.

In experiment 4A, the participants' d's at the verb-region suggest that they were already linking fillers with gaps correctly even in the iterative (–VOICE). However, they showed greater sensitivity when the verb exhibited AV (+VOICE). Thus, there was an added benefit to having voice, even when participants had already committed to an interpretation.

In experiment 4B, their d's suggest that only when the verb exhibited PV did they commit to an interpretation. At the co-argument-region, an asymmetric pattern—similar to what was observed with AV and iteratives at the verb-region—emerged: even though their d's suggest they were linking fillers with gaps correctly even in recent perfectives, they showed greater sensitivity when the verb exhibited PV. Thus, there was an added benefit to having voice, even when all of the information to complete the dependency had already been encountered.

There is also the question of why there was a plausibility effect at the ang-region. At this point of the presentation, the verb had yet to be seen, so the rejections could not possibly be attributed to plausibility. I argue that this is in fact an animacy effect. During debriefing and subsequent fieldwork, I learned that there are (at least) two types of *d*-linkers in Tagalog. Some speakers use *alin* 'which' as a catch-all *d*-linked interrogative, irrespective of the animacy of the noun being *d*-linked. On the other hand, some speakers have a restricted environment for the *d*-linked interrogative *alin*. These speakers can only use *alin* with inanimate nouns. In order for an animate noun to be *d*-linked, *sino* 'who' needs to be used. For example, these speakers would use *sinong dalaga* 'which young woman' instead of *aling dalaga*. In an SMS-task, this would mean that these speakers would judge sentences with *alin* followed by an animate noun as unacceptable and would choose to reject them before encountering the elements downstream. This is entirely consistent with the results observed in Experiment 4. In experiment 4A, plausible sentences (i.e., those with animate fillers) had higher rejection rates than implausible sentences (i.e., those with inanimate fillers). In experiment 4B, implausible sentences (i.e., those with animate fillers) had higher rejection rates than plausible sentences (i.e., those with inanimate fillers). In other words, sentences involving animate fillers had higher rejection rates than those involving inanimate fillers across sub-experiments.

To reiterate, experiment 4 establishes that voice is a cue that comprehenders attend to and use when processing FGDs. Even though Tagalog comprehenders actively associated the filler with the gap before encountering the co-argument, having voice on the verb enhanced this association.

4.4 Experiment 5: Comparing voice cues

It remains an open question as to whether comprehenders use the different types of voice the same way. Experiment 4 did not allow me to directly compare how they used

AV and PV because I used different baselines, ones that allowed me to ask the whether-question. It also remains an open question as to whether comprehenders use voice the same way across different types of dependencies. Thus, the goal of Experiment 5 is to directly compare the time courses of dependency formation when the verb exhibits AV and PV using three different types of FGDs. In other words, in Experiment 5, I ask the how-question.

4.4.1 Participants

I recruited 85 participants from UP–Diliman using the same recruitment criteria as in Experiments 1–4. They ranged from 18 to 44 years old (M = 25, SD = 6), and all lived in and around Metro Manila at the time of testing. They received a Starbucks gift certificate valued at 400 PHP for participating in the study.

4.4.2 Materials, procedure, and analysis

The experiment employed a 2×2 factorial design, crossing whether the verb exhibited AV or PV (VOICE: AV, PV) and whether the filler was globally plausible (PLAUSIBILITY: +PLAUS, −PLAUS). I created 3 sets of 24 items involving non-reversible predicates in *wh*-questions, RCs, and *ay*-inverted sentences. The experimental items were combined together in one session and randomized with another set of 24 items involving run-of-the-mill declarative sentences. Each item was distributed evenly via Latin square design. The same 24 verbs were used throughout and thus, they appeared four times per participant in different constructions and with different co-argument pairs.

To discourage participants from singling out *wh*-questions as the only non-declarative clause-type, I used embedded questions. To discourage them from creating expectations about when to provide a response, I varied the region where the dependency started: for *wh*-questions, the filler was in region 4; for RCs, it was in region 6; and for *ay*-inverted

sentences, it was in region 2. Crucially, the distance between the filler, the verb, and the co-argument remained consistent across dependencies. Table 4.2 provides a sample of the experimental items using *bili* 'buy.'

Experiment 5 used the same procedure as Experiment 4. I also used the same dependent measure as in experiment 4 to compare the differences in the time courses of sentences with AV and those with PV when interpreting three types constructions. Their cumulative *d'*s were derived using the same bootstrapping procedure.

4.4.3 Results and discussion

The main findings of experiment 5 are the following:

- Tagalog comprehenders actively associated the filler with the gap even before the fully disambiguating co-argument;
- The way they used voice may be mediated by voice-type and FGD-type.

Reported in Figure 4.3 are the participants' phrase-by-phrase mean cumulative *d'*s and the 95% CI by VOICE at each region.

Experiment 5A: Comparing AV and PV in *wh*-questions

There were no plausibility effects observed at the filler-region. At the *ang*-region, participants were rejecting implausible sentences more than plausible ones when the verb exhibited PV, but not when it exhibited AV: d' = -.33, 95% CI [-.73, .14] for AV and d' = .60, 95% CI [.23, .86] for PV. This is further corroborated by a significant difference in *d'*s: $\Delta d'$ = -.94, 95% CI [-1.46, -.36].

Table 4.2: Sample experimental items used in experiment 5, manipulating VOICE: AV, PV and PLAUSIBILITY: +, −.

Exp. 5A: WHQ				Region			
VOICE	PLAUS	Filler	ang	Verb	XP	CoArg	...
AV	+PLAUS	*aling babae*	*ang*	*bumibili*	*gabi-gabi*	*ng alak*	*sa Trinoma*
		which woman		buy.AV	every night	wine	at Trinoma ...
		(Ernesto knows) which woman buys wine at Trinoma every night ...					
AV	−PLAUS	*aling alak*	*ang*	*bumibili*	*gabi-gabi*	*ng babae*	*sa Trinoma*
		which wine		buy.AV	every night	woman	at Trinoma ...
		#(Ernesto knows) which wine buys the woman at Trinoma every night ...					
PV	+PLAUS	*aling alak*	*ang*	*binibili*	*gabi-gabi*	*ng babae*	*sa Trinoma*
		which wine		buy.PV	every night	woman	at Trinoma ...
		(Ernesto knows) which wine the woman buys at Trinoma every night ...					
PV	−PLAUS	*aling babae*	*ang*	*binibili*	*gabi-gabi*	*ng alak*	*sa Trinoma*
		which woman		buy.PV	every night	wine	at Trinoma ...
		#(Ernesto knows) which woman the wine buys at Trinoma every night ...					

Exp. 5B: RC				Region			
VOICE	PLAUS	Filler	na	Verb	XP	CoArg	...
AV	+PLAUS	*tatay*	*na*	*bumibili*	*taon-taon*	*ng kotse*	*sa Toyota*
		father		buy.AV	every year	car	at Toyota ...
		(Rafael thought) the father who buys a car every year at Toyota... (is very small)					
AV	−PLAUS	*kotse*	*na*	*bumibili*	*taon-taon*	*ng tatay*	*sa Toyota*
		car		buy.AV	every year	father	at Toyota ...
		#(Rafael thought) the car that buys a father every year at Toyota... (is very small)					
PV	+PLAUS	*kotse*	*na*	*binibili*	*taon-taon*	*ng tatay*	*sa Toyota*
		car		buy.PV	every year	father	at Toyota ...
		(Rafael thought) the car that the father buys every year at Toyota... (is very small).					
PV	−PLAUS	*tatay*	*na*	*binibili*	*taon-taon*	*ng kotse*	*sa Toyota*
		father		buy.PV	every year	car	at Toyota ...
		#(Rafael thought) the father that the car buys every year at Toyota... (is very small)					

Exp. 5C: AY		Region					
VOICE	PLAUS	Filler	ay	Verb	XP	CoArg	...
AV	+PLAUS	*ang nanay*	*ay*	*bumibili*	*sa KFC*	*ng pagkain*	*kasi*
		mother		buy.AV	at KFC	food	because ...
		(Every day,) the mother, she buys food at KFC because (she has no time...)					
AV	−PLAUS	*ang pagkain*	*ay*	*bumibili*	*sa KFC*	*ng nanay*	*kasi*
		food		buy.AV	at KFC	mother	because ...
		#(Every day,) the food, it buys a mother at KFC because (she has no time...)					
PV	+PLAUS	*ang pagkain*	*ay*	*binibili*	*sa KFC*	*ng nanay*	*kasi*
		food		buy.PV	at KFC	mother	because ...
		(Every day,) the food, the mother buys it at KFC because (she has no time...)					
PV	−PLAUS	*ang nanay*	*ay*	*binibili*	*sa KFC*	*ng pagkain*	*kasi*
		mother		buy.PV	at KFC	food	because ...
		#(Every day,) the mother, the food buys her at KFC because (she has no time...)					

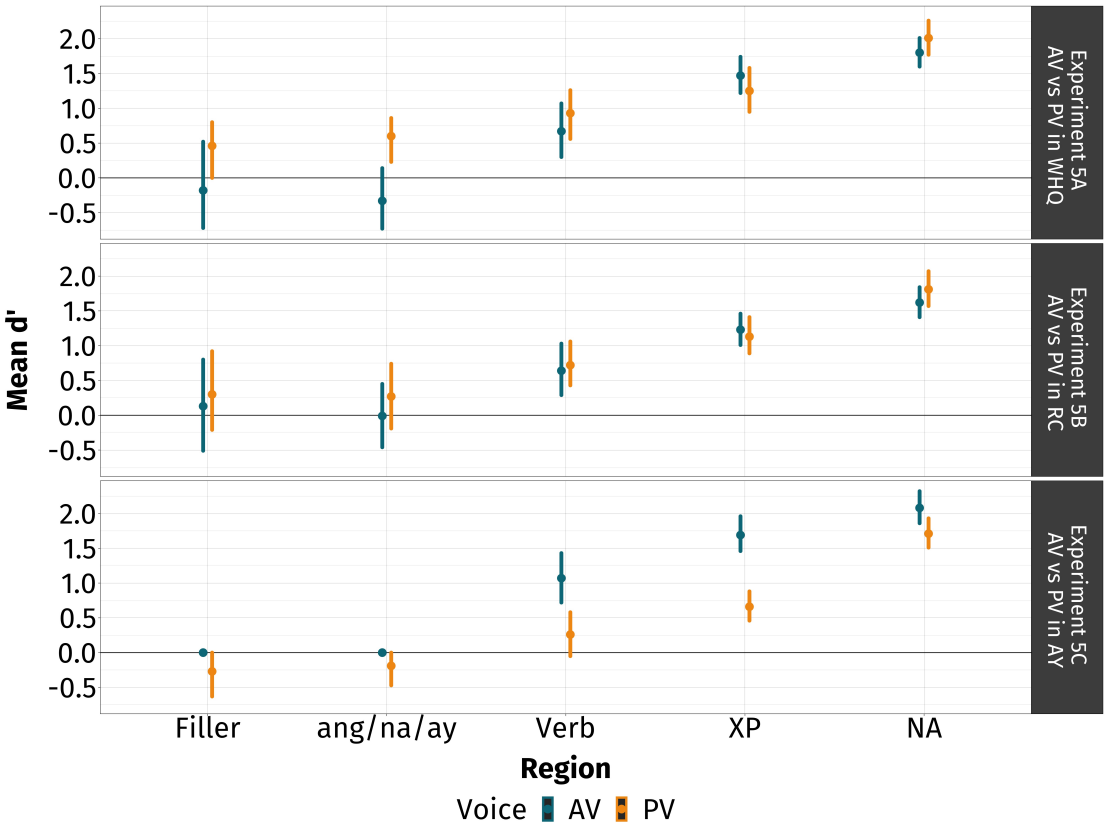


Figure 4.3: Phrase-by-phrase discrimination scores in Experiment 5. Plotted are the mean scores at each region (points) with 95% confidence intervals (bands) derived by a bootstrap over participants. In blue are AV-conditions and in gold are PV-conditions. The topmost panel represents the comparison between AV and PV in *wh*-questions (experiment 5A), the middle panel represents the comparison between AV and PV in relative clauses (experiment 5B), and the bottommost panel represents the comparison between AV and PV in *ay*-inverted sentences (experiment 5C). The verb- and XP-regions, and the co-argument-region are the critical and disambiguating regions, respectively.

At the verb-region, they were rejecting implausible sentences more than plausible ones: d' = .67, 95% CI [.30, 1.07] for AV and d' = .93, 95% CI [.56, 1.26]. Their sensitivity was comparable across different voices: $\Delta d'$ = -.26, 95% CI [-.73, .25]. Their rejection behavior at the verb-region persisted into the XP- and co-argument-regions. They were rejecting implausible sentences more than plausible ones, and their sensitivity was comparable across different voices.

Experiment 5B: Comparing AV and PV in relative clauses

There were no plausibility effects in the filler- and *na*-regions. At the verb-region, participants were rejecting implausible sentences more than plausible ones: d' = .64, 95% CI [.29, 1.03] for AV and d' = .72, 95% CI [.43, 1.06] for PV. Their sensitivity was comparable across different voices: $\Delta d'$ = -.08, 95% CI [-.59, .42]. Their rejection behavior at the verb-region persisted into the XP- and co-argument-regions. They were rejecting implausible sentences more than plausible ones, and their sensitivity was comparable across different voices.

Experiment 5C: Comparing AV and PV in *ay*-inverted sentences

There were no plausibility effects in the filler- and *ay*-regions. At the verb-region, we found an asymmetry in how participants used voice. They were rejecting implausible sentences more than plausible ones when the verb exhibited AV, but this was not the case when the verb exhibited PV: d' = 1.07, 95% CI [.72, 1.43] for AV and d' = .26, 95% CI [-.05, .58].

At the XP-region, they were rejecting implausible sentences more than plausible ones across the different types of voice: d' = 1.69, 95% CI [1.46, 1.96] for AV and d' = .66, 95% CI [.46, .88] for PV. The asymmetry found at the verb-region persisted into the XP-region: $\Delta d'$ = 1.03, 95% CI [.70, 1.37].

At the co-argument-region, they were rejecting implausible sentences more than plausible ones across the different types of voice: d' = 2.08, 95% CI [1.86, 2.32] for AV and d' = 1.71, 95% CI [1.51, 1.93] for PV. Their sensitivity was comparable across different voices: $\Delta d'$ = .38, 95% CI [.13, .62].

4.4.4 Discussion

The goal of experiment 5 was to directly compare the time courses of dependency formation when the verb exhibited AV and PV in three different types of FGDs. The main findings are the following. First, comprehenders actively associated the filler with the gap even before the fully disambiguating co-argument. This replicates the findings of Experiment 4. Second, the way in which voice was used varied across different voice types and across different dependencies.

The participants' *d'*s at the verb- and XP-regions provide evidence that voice was used a cue when interpreting FGDs. Participants were correctly rejecting implausible sentences as early as the verb- and XP-regions, even before they encountered the disambiguating co-argument. These findings replicated Experiment 4, and are consistent with studies supporting active dependency formation (Crain & Fodor, 1985; L. Frazier, 1987; L. Frazier & Clifton, 1987, *inter alia*). They are also consistent with the idea that comprehenders use various sources of linguistic information—lexico-semantic (Boland et al., 1990, 1995; Trueswell et al., 1994) and morphosyntactic ones (Kamide et al., 2003; Wagers et al., 2015), for example—to form said dependency.

In *wh*-questions and RCs, comprehenders used AV and PV symmetrically based on their comparable *d'*s. As early as the verb, voice—both AV and PV—had allowed comprehenders to already commit to the correct interpretation. In contrast, in *ay*-inverted sentences, comprehenders exhibited an asymmetry. Their commitment to the correct interpretation was much earlier when the verb exhibited AV than when it exhibited PV. Their *d'* for AV was significantly positive at the verb-region. Their *d'* for PV, on the other hand, was only significantly positive at the XP-region—and even then, they still showed greater sensitivity when the verb had AV.

In *wh*-questions, comprehenders were also rejecting implausible sentences more than plausible ones when the verb exhibited PV at the ang-region. I argue that this is an ani-

macy effect, much like what was found in Experiment 4. The *d*-linked interrogative *alin* is compatible with inanimate fillers, but not animate ones. In an SMS-task, this would mean that these speakers would judge sentences with *alin* followed by an animate noun as unacceptable. This is consistent with the effect that we found for PV at the *ang*-region.

Why must voice be used differently across different FGDs? One straightforward hypothesis is that voice is not equally as informative with regard to the mapping between voice and the identity of the gap, contra the hypothesis presented in Section 4.2, in the dependencies investigated. The results of Experiment 1 indicated that AV-morphology was a strong diagnostic of a subject gap, while PV-morphology was less diagnostic of a gap corresponding to the patient role. Thus, in a sense, the results of Experiment 5 are consistent with the results of Experiment 1.

4.5 General discussion

Tagalog comprehenders used voice to reliably link the filler to the gap and commit to the correct interpretation before encountering disambiguating information in the form of the co-argument. Thus, I add Tagalog to the general picture that comprehenders actively link the filler to the gap and they can use morphological information, like voice morphology, to do so. However, there was a wrinkle: the way in which comprehenders leveraged voice as a cue was mediated by the type of FGDs.

In what follows, I first discuss the level of granularity of the information about voice that comprehenders use in real-time by relating the results of Experiments 1 and 5. I discuss reasons why RCs patterned like *wh*-questions, and why they did not pattern like *ay*-inverted sentences. I then look at the asymmetries found in Experiment 5 and view them as a window to what other classes of information could interact when comprehenders process FGDs. When Tagalog comprehenders incrementally interpret these dependencies, they engage with different classes of linguistic information that overlap in time and

priority, including (but not limited to) information about the structural similarities or the thematic complexity of the dependencies involved, the relative frequency of the different types of voice in the language, and finer-grained information like how informative voice is as a cue in a given FGD.

4.5.1 Tagalog comprehenders attend to voice information at the FGD-level

We know from experiment 4 that comprehenders are able to use morphological information. The level of granularity that comprehenders actually use in real-time, however, remains an open question. Here I discuss how the results of Experiment 1 can shed some light on the results of Experiment 5.

The hypothesis presented in the beginning of Section 4.2 of how voice can be used in FGD-processing operates at the grammar-level, aggregating over the intimate correlations between voice morphology and the extracted nominal in the many FGDs encountered in the input. In other words, this could be encoded as coarse-grained information about voice morphology. If comprehenders rely on voice at this level, their behavior in an SMS task should be highly symmetrical. In other words, AV and PV should be leveraged similarly across different types of dependencies since they are a strong diagnostic of a gap corresponding to the subject/agent and the object/patient, respectively.

On the other hand, if comprehenders rely on finer-grained information about voice, then their behavior in an SMS task might mirror the results of Experiment 1 more. AV was a highly informative cue irrespective of the type of dependency and thus, should be a robust trigger for dependency formation in *all* environments. Meanwhile, PV was variably informative, depending on the particular type of FGD. Thus, PV should trigger less predictive parsing, as it becomes less diagnostic of a gap corresponding to the object/patient role. More specifically, PV should be the most predictive in *wh*-questions and

least predictive in *ay*-inverted sentences and RCs.

The participants' symmetrical discrimination behavior in *wh*-questions (Experiment 5A) is consistent with comprehenders relying on either coarse-grained or finer-grained information. Under a view where the use of voice morphology operates at the grammar-level, comprehenders used AV and PV symmetrically by hypothesis. Under a view where the use of voice morphology attends to the different levels of dependency-specific informativity, comprehenders used AV and PV symmetrically because both voices are highly informative of the mapping between voice and the identity of the gap. Upon encountering the verb, AV and PV were good diagnostics of gaps corresponding to the agent and the patient, respectively.

When we consider their behavior with regard to the other two FGDs, the predictions of the different levels of granularity of voice are pulled in opposite directions, as seen in Table 4.5.

Table 4.5: A summary of their discrimination behavior in experiment 5 by the different levels of granularity of voice morphology. A "✓" represents consistency of the results with the predictions for that level of granularity. A "✗" represents inconsistency of the results with the predictions for that level of granularity.

	WHQ	RC	AY
General (coarse-grained)	✓	✓	✗
FGD-specific (fine-grained)	✓	✗	✓

First, comprehenders used AV and PV symmetrically in RCs (Experiment 5B). That is, they were reliably committing to the correct interpretation as early as the verb-region. This symmetrical behavior is consistent with the view where voice, by hypothesis, operates at the grammar level. However, this is inconsistent with the view that comprehenders attend to the information encoded by voice at the FGD-specific level. If they did, comprehenders, when the verb has PV, should be the least predictive in RCs. We saw in Experiment 1 that voice was the least informative of the mapping between voice and the

identity of the gap in RCs.

Second, comprehenders used AV and PV asymmetrically in *ay*-inverted sentences (Experiment 5C). That is, AV allowed them to reliably commit to the correct interpretation as early as the verb-region, while PV only allowed them to reliably commit to the correct interpretation as early as the XP-region. This asymmetrical behavior is inconsistent with the view where voice operates at the grammar level. However, this is consistent with the view that comprehenders attend to finer grained information about voice. In Experiment 1, AV was more informative than PV in this dependency. It is unsurprising then that participants exhibited earlier sensitivity to implausibility when the verb had AV compared to when it had PV.

I argue that comprehenders attended to FGD-specific information about voice. Under this view, their behavior in RCs becomes curious. We would expect comprehenders to use voice asymmetrically in a way that closely resembled their discrimination behavior in *ay*-inverted sentences at the very least. However, their discrimination behavior in Experiment 5B suggested that they used voice symmetrically in a way that more closely resembled their behavior in *wh*-questions. Even though comprehenders ultimately used AV and PV symmetrically, there is evidence for an asymmetry, comparable to the one seen in *ay*-inverted sentences, in what I refer to as the participants' UNREJECTED READING TIMES. These are the reading times of the trials where the participant chose to continue reading. I discuss this in greater detail in the subsection below.

Crucial here is that comprehenders in Experiment 5B exhibited a greater plausibility mismatch effect when the verb exhibited AV at the XP- and co-argument-regions. Their unrejected reading times were longer in the implausible conditions compared to their reading times in the plausible conditions, especially when the verb exhibited AV. I take this to suggest that participants were sensitive to the plausibility mismatch even if they had decided to delay the rejection of an implausible item at these regions. This finding

is consistent with the claim that comprehenders are paid attention to how informative voice is at the FGD-level, but their ability to use this level of information in real-time could be modulated by other classes of information, like structural similarity or thematic complexity, for example.

4.5.2 Unrejected reading times: A more nuanced picture of choice

The main dependent measure used in the analyses above was the participants' *d'*, which was calculated by scaling the participants' correct rejections of implausible items against their their incorrect rejections of plausible items (Macmillan & Creelman, 2005). There is a secondary measure that can be used to have a more nuanced picture of the participants' choice to accept or reject: their unrejected reading times. These are the reading times of the trials where they chose to continue reading. This measure is akin to the dependent measure used in traditional self-paced reading experiments.

For any given trial, at each region, participants had the option to either reject or continue reading. On the trials where they chose to continue reading, their unrejected reading time for that region was recorded. I used their unrejected RTs as a secondary measure to determine whether participants were sensitive to plausibility mismatches even if they decided to delay the rejection of implausible items. I analyzed their log-transformed RTs at two regions of interest—the Verb- and CoArg-regions—using *lme4* (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2018). The fixed effects were VOICE (+VOICE, −VOICE), PLAUSIBILITY (+PLAUS, −PLAUS) and their interaction. These factors were sum-coded, such that +VOICE and +PLAUS mapped to the negative coefficients. I included the maximal random effects structure that allowed the models to converge (Barr, 2013; Barr et al., 2011). The *p*-values reported were generated using the Satterthwaite's method in the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017).

One key difference between cumulative *d'* and unrejected RTs is the degree of com-

mitment to an interpretation that a comprehender must have given what she has encountered so far. When she chose to reject at a region in any given trial, she had committed to a particular interpretation: what continuations were possible and what continuations were not. In contrast, when she chose to continue reading but displayed slower readings times, she might have experienced processing difficulty—due to a plausibility mismatch effect, for example—but she may have been unable or unwilling to fully commit to an interpretation just yet (Mauner & Koenig, 2000). Together, these measures provide a more nuanced picture of the time courses of FGD-processing.

In Table 4.6 are the mean unrejected reading times in *ms* by VOICE and PLAUSIBILITY at the critical and disambiguating regions (i.e., the verb-, XP-, and co-argument-regions). In Table 4.7 are the estimated models. In the text that follows, I only report significant effects on their unrejected RTs.

Experiment 5A: Comparing AV and PV in *wh*-questions

At the verb-region, their unrejected RTs indicated that on trials where they chose to continue reading, they took longer when the verb exhibited PV (M = 830 *ms*) than when it exhibited AV (M = 781 *ms*). This effect was qualified by a significant interaction, such that they read implausible sentences significantly slower than plausible ones when it exhibited PV (ΔRT = 89 *ms*) than when it exhibited AV (ΔRT = -56 *ms*).

At the XP-region, their RTs were longer in the implausible conditions (M = 1168 *ms*) compared to the plausible conditions (M = 978 *ms*). At the co-argument-region, their RTs were again longer in the implausible conditions (M = 1343 *ms*) compared to the plausible conditions (M = 969 *ms*). This effect was qualified by a significant interaction, such that the difference between plausible and implausible sentences was greater when the verb exhibited PV (ΔRT = 447 *ms*) than when it exhibited AV (ΔRT = 299 *ms*). They were significantly faster when reading plausible sentences with PV.

Table 4.6: Mean unrejected reading times (*ms*) at the Verb-, XP-, and co-argument-regions by VOICE and PLAUSIBILITY with the standard error of the mean (*SE*) and the difference in reading times (ΔRT: $RT_{-VOICE} - RT_{+VOICE}$)

			Verb-region			XP-region			Co-Arg-region		
	VOICE	PLAUS	*M*	*SE*	ΔRT	*M*	*SE*	ΔRT	*M*	*SE*	ΔRT
Exp. 5A WHQ	AV	+	808	28	-56	1019	37	125	1037	39	299
	AV	–	752	25		1144	47		1336	77	
	PV	+	787	42	89	938	32	248	902	24	447
	PV	–	876	31		1186	44		1349	76	
Exp. 5B RC	AV	+	729	18	64	876	33	222	948	35	291
	AV	–	793	33		1098	47		1239	76	
	PV	+	768	33	239	928	41	78	1009	42	125
	PV	–	1007	177		1006	47		1134	62	
Exp. 5C AY	AV	+	745	33	19	953	37	1	1214	66	240
	AV	–	764	32		954	42		1454	116	
	PV	+	786	28	90	957	55	107	1122	38	225
	PV	–	876	44		1064	41		1347	68	

Table 4.7: Summary of linear mixed-effects models estimated in Experiment 5. The models included the log-transformed unrejected RTs as the dependent measure, and sum-coded VOICE, PLAUSIBILITY and their interaction into the models as fixed effects. Random effects included the maximal structure that converged.

	Verb-region			XP-Region			Co-Arg-Region		
	Estimate	*SE*	*t*	Estimate	*SE*	*t*	Estimate	*SE*	*t*
Exp. 5A: *Wh*-questions									
Intercept	6.53	.03	244.26**	6.81	.05	126.98**	6.86	.05	143.69**
VOICE	.04	.02	2.02*	-.006	.02	-.27	-.03	.03	-1.05
PLAUS	.03	.02	1.36	.18	.03	5.83**	.25	.05	5.43**
VOICE×PLAUS	.02	.04	3.86**	.08	.04	1.87†	.11	.05	2.00*
Exp. 5B: Relative clauses									
Intercept	6.51	.03	197.94**	6.72	.05	131.20**	6.80	.04	192.58**
VOICE	.02	.02	.80	-.05	.03	-1.57	-.02	.05	-.54
PLAUS	.06	.02	2.64**	.19	.03	5.55**	.19	.03	6.08**
VOICE×PLAUS	.04	.04	.91	-.15	.05	-3.33**	-.12	.06	-2.01*
Exp. 5C: *Ay*-inverted sentences									
Intercept	6.50	.03	192.04**	6.72	.05	141.46**	6.92	.06	124.08**
VOICE	.07	.02	3.26*	.02	.03	.74	-.001	.04	-.03
PLAUS	.04	.02	1.95†	.01	.02	4.61**	.17	.05	3.61*
VOICE×PLAUS	.01	.04	.34	.05	.05	1.02	-.004	.06	-.06

$^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .001$

Experiment 5B: Comparing AV and PV in relative clauses

At the verb-region, their unrejected RTs suggest that on trials where they chose to continue reading, they still detected the plausibility mismatches. Their RTs were longer in the implausible conditions (M = 902 *ms*) compared to the plausible ones (M = 748 *ms*).

At the XP-region, their RTs were again longer in the implausible conditions (M = 1047 *ms*) than in the plausible ones (M = 902 *ms*). This effect was qualified by a significant interaction, such that the difference between plausible and implausible sentences was greater when the verb exhibited AV (ΔRT = 222 *ms*) than when it exhibited PV (ΔRT = 78 *ms*).

At the co-argument-region, their RTs were again longer in the implausible conditions (M = 1180 *ms*) compared to the plausible ones (M = 980 *ms*). This effect was again qualified by a significant interaction, such that the difference between plausible and implausible sentences was greater when the verb exhibited AV (ΔRT = 291 *ms*) than when it exhibited PV (ΔRT = 125 *ms*).

Experiment 5C: Comparing AV and PV in *ay*-inverted sentences

At the verb-region, their unrejected RTs suggest that on trials where they chose to continue reading, they took longer when the verb exhibited PV (M = 831 *ms*) compared to when it exhibited AV (M = 754 *ms*).

At the XP-region, their RTs were longer in the implausible conditions (M = 1017 *ms*) compared to the plausible ones (M = 955 *ms*).

At the co-argument region, their RTs were longer in the implausible conditions (M = 1392 *ms*) compared to the plausible ones (M = 1168 *ms*).

Discussion

Recall that the goal of Experiment 5 was to directly compare the time courses of dependency formation when the verb exhibited AV and PV in three different types of filler-

gap dependencies. I found the following. First, I replicated the finding from Experiment 4 that comprehenders actively associated the filler with the gap even before the fully disambiguating co-argument. Second, I found that the way in which voice was used varied across different voice types and across different dependencies.

As we saw, the participants' *d'*s at the verb- and XP-regions provide evidence that voice was used a cue when interpreting FGDs. Participants were correctly rejecting implausible sentences as early as the verb- and XP-regions, even before they encountered the disambiguating co-argument. Their unrejected RTs provided further evidence that they were using voice as a cue. On trials where participants decided to continue reading, their longer reading times for implausible sentences suggest that they did detect the implausibility of the sentence even before the co-argument was encountered.

When their unrejected RTs are considered, a more nuanced picture emerges. There was a PV-advantage at the co-argument-region in *wh*-questions and an AV-advantage at the XP-region in relative clauses. I address these in the next section, where I talk about asymmetries in greater detail.

4.5.3 (A)symmetries as a window to interactions

Comprehenders can prioritize different types of cues (Tamaoka, Sakai, Kawahara, Lim, & Koizumi, 2005) and they can even ignore very reliable cues in favor of less reliable ones (Gagliardi & Lidz, 2014; MacWhinney et al., 1984). Here, I look at how finer-grained information about voice interacts with other classes of linguistic information available to the comprehender. Below I discuss the apparent (a)symmetries found in Experiment 5 and use them as windows to these interactions.

There are (at least) two ways of accounting for the curious case of the participants' behavior in RCs. The first way is to offer an explanation as to why RCs patterned like *wh*-questions in Experiment 5—despite having very different levels of informativity of

the mapping between the particular form of voice and the identity of the gap, as operationalized in Experiment 1. Structural similarity could be a reason why RCs patterned like *wh*-questions in Experiment 5 despite the results of Experiment 1.

One construal of structural similarity is the relatedness of the constructions involved. As discussed in Chapter 2, argument *wh*-questions and RCs are structurally similar in Tagalog, since argument *wh*-questions are essentially RCs (Aldridge, 2002, 2013; Kroeger, 1993). On the other hand, *ay*-inverted sentences are not related to the other two. Another construal of structural similarity is the similarity in clausal complexity, defined here as the number of clauses involved in the dependency. Both *wh*-questions and RCs involve bi-clausal structures and thus have comparable clausal complexities. On the other hand, *ay*-inverted sentences only involve mono-clausal structures (Kroeger, 1993; Sabbagh, 2014). Under either construal of structural similarity, the structural similarity of *wh*-questions and RCs was prioritized over the information about the mapping between voice and the identity of the gap, and thus, structural similarity exerted a more powerful influence on the participants' response behavior.

The second way of accounting for the curious case of RCs is to offer an explanation as to why RCs did not pattern like *ay*-inverted sentences in Experiment 3—despite having very comparable level of informativity of the mapping between the particular form of voice and the identity of the gap, as operationalized in Experiment 1. Thematic complexity could be a reason why RCs did not pattern like *ay*-inverted sentences despite the results of Experiment 1. As mentioned before, the filler in RCs serves one thematic role in the matrix clause and another thematic role in the embedded clause. In contrast, the filler in *ay*-inverted sentences serves only one thematic role. Under this view, the thematic requirements of the filler was prioritized over the information about the mapping between voice and the identity of the gap, and thus exerted a more powerful influence on the participants' response behavior.

In *wh*-questions, even though comprehenders ultimately used AV and PV symmetrically, there is evidence for an asymmetry in their unrejected reading times. At the co-argument-region, their unrejected reading times were longer in the implausible conditions compared to their reading times in the plausible conditions, especially when the verb exhibited PV. In other words, upon encountering the disambiguating co-argument, comprehenders were more sensitive to plausibility mismatches when the verb exhibited PV. This difference cannot be attributed to the varying levels of informativity of the mapping between voice and the identity of the gap, since AV and PV are equally informative in *wh*-questions. Thus, something else must be underlying this effect. One potential explanation for this is the relative frequency of AV and PV in the language. Corpus studies have found that PV is more frequent than AV in transitive sentences (Cooreman et al., 1984; Garcia, Dery, Roeser, & Höhle, 2018). One way to reframe this difference in relative frequency is that comprehenders have more experience integrating incoming information when the verb has PV. Crucially, the plausibility mismatch effect found at the co-argument restriction is consistent with the view that comprehenders prioritized the information about the mapping between voice and the identity of the gap over the information about the relative frequencies of voice.

4.5.4 Issues and future directions

First, it would have been ideal to perform corpus-based analyses to quantitatively derive the mapping between voice and the identity of the gap, in the same vein as the first study of Garcia et al. (2019) but with FGDs. However, for practicality reasons, I used the results of Experiment 1 to estimate how informative voice is as a cue. In order to use a corpus to calculate the informativity of voice as a cue to detect when the extracted nominal is not cross-referenced by the voice on the verb, the corpus would need to be part-of-speech tagged, as well as be dependency parsed. I do not know of any corpora

that satisfy these requirements.

Second, I related these judgment-estimated values to their behavior in an SMS task. Given the obvious differences between the two methodologies, it is perhaps unsurprising if these two tasks elicited different ways voice is to be leveraged as a cue to guide interpretation. Perhaps their behavior in the SMS task was a reflection of how voice could be used when comprehenders were forced into a strictly incremental mode of parsing with no option to backtrack. On the other hand, their behavior in an untimed judgment study was a reflection of how voice could be used to make terminal state judgments, that is, when all of the bottom-up information became available to the comprehender. I fully acknowledge that using the judgment study was an imperfect proxy for estimating how informative voice is as a cue. I have no way of completely ruling out this possibility.

Third, it is difficult to adjudicate between the two construals of structural similarity, given the nature of the experimental items. In future studies, I could control for the number of clauses involved in the dependencies and isolate the effect of the relatedness of the constructions on the participants' behavior by investigating the participants' behavior in long-distance *ay*-inversion. In (11-a) is the basic word order of a sentence involving the verb *kailangan* 'need.' It takes a complement clause, demarcated with "[ ]," and inside that complement clause, the nominative argument is in its canonical position. That is, there is no extraction involved. In (11-b) is an instance where *sina Ben* undergoes local *ay*-inversion. My experimental items are akin to this example, modulo the embedding context. Finally, in (11-c) is an instance where *sina Ben* undergoes long-distance *ay*-inversion.

(11) Adapted from Kroeger (1993, p. 125)

a. Kailangan-g [mag-tanim ng gulay sina Ben]
need-that AV-plant GEN vegetable NOM.PL PN
'It is necessary for Ben and the others to plant vegetables.'

b. Kailangan-g [sina Ben ay mag-tanim ng gulay]
need-that NOM.PL PN AY AV-plant GEN vegetable

'As for Ben and the others, it is necessary for them to plant vegetables.'

c. Sina Ben ay kailangan-g [mag-tanim ng gulay]
NOM.PL PN TOP need-that AV-plant GEN vegetable
'As for Ben and the others, it is necessary for them to plant vegetables.'

The configuration of *ay*-inversion in (11-c) now matches the clausal complexity of *wh*-questions and RCs in Experiment 5. This allows us to partial out the contribution of clausal complexity to the participants' response behavior. If clausal complexity was the main driving force of why RCs and *wh*-questions patterned together in experiment 5, then we expect to see long-distance *ay*-inversion to pattern like the other two now that the number of clauses is held constant across FGDs. On the other hand, if the relatedness of the constructions was the driving force of why RCs and *wh*-questions, then we expect to see the same general pattern. RCs and *wh*-questions would still pattern together, to the exclusion of long distance *ay*-inversion.

Similarly, long-distance *ay*-inversion can also be used to see how tenable of an explanation thematic complexity is for explaining why RCs did not pattern like *ay*-inverted sentences. Given the nature of the experimental items, the contributions of clausal complexity and thematic complexity on why RCs did not pattern like *ay*-inverted sentences were confounded. Were they different because RCs involved bi-clausal structures, while *ay*-inverted sentences involved mono-clausal structures? Or were they different because RCs assign two roles to the filler, while *ay*-inverted sentences only assign one role? The configuration of *ay*-inverted sentences in (11-c) now matches the clausal complexity of RCs, but crucially, no new thematic relations are introduced.

Lastly, AV's higher level of informativity is compatible with what has been called the "subject preference" in the processing literature (Bickel, Witzlack-Makarevich, Choudhary, Schlesewsky, & Bornkessel-Schlesewsky, 2015). There is a preference to analyze what is initially an ambiguous NP as being a subject, a cover term to refer to the sole

argument of an intransitive verb or the agent of a transitive verb. In a sense, the current findings also point in a similar direction. Agents in Tagalog are somehow privileged in filler-gap dependency processing, as exemplified by the AV-PV asymmetry in *ay*-inverted sentences—and potentially, RCs, as well. A similar asymmetry has been observed by Tanaka et al. (2019) in Tagalog-speaking children's acquisition of relative clauses, and by Bondoc et al. (2018) in the development of the relativization grammar of aphasics. Sauppe (2016) also observed that agents are privileged when processing canonical declaratives in the language. The behavior of intransitive subjects remain unknown because the studies so far—the present chapter included—have focused only on predicates with two arguments.

Note that AV's higher level of informativity is equally compatible with what has been called the "agent preference" in the literature (Cohn & Paczynski, 2013). Because agent preference is a proper subset of the subject preference, it is difficult to ascertain whether it can be attributed to subjecthood or agenthood. In future studies, I could investigate the comprehenders' behavior when processing two different types of intransitives (Perlmutter, 1978): unergatives, whose sole argument is an agent (e.g., *walk*), and unaccusatives, whose sole argument is a patient (e.g., *sleep*). In doing so, I can investigate which preference could be operative in Tagalog and could be a potential explanation for the higher level of informativity that AV has compared to PV. If the advantage is potentially due to subjecthood, then we would expect unergatives and unaccusatives to pattern together. On the other hand, if the advantage is potentially due to agenthood, then we would expect unergatives to exhibit an advantage over unaccusatives in processing.

Chapter 5

The role of word order in relative clause processing

The main goal of this chapter is to investigate how relative clauses are processed in Tagalog and see what classes of information Tagalog comprehenders might be paying attention to when processing these types of dependencies. RCs have been—and will most likely continue to be—a staple in psycholinguistics, language acquisition, and formal linguistics. The reason why language researchers have been "preoccupied" with investigating RCs is best summarized by Wagers et al. (2018, p. 207)

> RCs provide an unlimited number of ways for the reference of an NP to be narrowed, because they permit the NP to be modified by any property expressible by a sentence containing a variable. To achieve this expressiveness, there must be a mechanism for linking the NP to a position inside the RC—the position corresponding to the variable. RCs are clauses themselves and so form their own grammatical microcosm. And this is the source of complexity that has preoccupied so many investigations: the link between the NP and its site of interpretation cannot be formed without interacting with the many processes that combine to 'build' a clause.

As we will see below, one of the common findings in RC processing research is that comprehenders prefer RELATIVE CLAUSES WITH SUBJECT-GAPS (SRCs), as in (1-a), over RELA-

TIVE CLAUSES WITH OBJECT-GAPS (ORCs), as in (1-b). Throughout, I refer to this preference as the SUBJECT-OBJECT ASYMMETRY, or simply the asymmetry.

(1) Adapted from Kwon, Gordon, Lee, Kluender, and Polinsky (2010)

a. The reporter that ___ attacked the senator disliked the editor *SRC*

b. The reporter that the senator attacked ___ disliked the editor *ORC*

In this chapter, I ask whether (i) SRCs are easier to process than ORCs in Tagalog; and if they are, (ii) whether other factors can modulate this asymmetry. This chapter also asks what the empirical landscape of RC processing in Tagalog can tell us about the classes of explanations that have been proposed in the literature to account for what seems to be a robust asymmetry cross-linguistically.

Previous studies suggest that SRCs are easier to process than ORCs in Tagalog (Pizarro-Guevara, 2014; Bondoc et al., 2018; Tanaka et al., 2019). These studies, however, have only investigated head-initial relative clauses. It remains an empirical question as to whether the same holds true for head-final relative clauses. The experiments presented here are in a sense, a replication and an extension of these studies.

The chapter is organized as follows. Section 5.1 provides an overview of the well-known subject-object asymmetry and its cross-linguistic and cross-methodological robustness. Section 5.2 reviews the classes of explanations that have been proposed in the literature to account for the asymmetry. Section 5.3, on the other hand, reviews some of the factors that have been reported to reduce—and sometimes, even reverse—the asymmetry. Section 5.4 gives a bird's eye view of the experiments, which are then presented in Sections 5.5–5.7. Finally, Section 5.8 concludes by synthesizing the results and addressing what these can tell us about the classes of explanations in the literature to account for the asymmetry.

5.1 Subject-object asymmetry in relative clauses

One of the common findings in RC processing research is that comprehenders prefer SRCs over ORCs. This asymmetry comes in two flavors. First, in languages with unambiguous relative clauses, like English, for example, ORCs have been found to be more difficult to process than SRCs. In other words, sentences like (2-b) are found to be harder to process compared to sentences like (2-a).

(2) Adapted from Kwon et al. (2010)

a. The reporter that ___ attacked the senator disliked the editor *SRC*

b. The reporter that the senator attacked ___ disliked the editor *ORC*

What "harder" actually means varies from study to study, largely depending on the methodology that researchers have used. This could mean lower percent correct in a repetition task (Diessel & Tomasello, 2005), a picture-selection task (Tanaka et al., 2019), or a memory-load experiment (P. C. Gordon, Hendrick, & Levine, 2002); slower reading times in self-paced reading (Gibson, Desmet, Grodner, Watson, & Ko, 2005); or in eye-tracking while reading, higher incidence of regressive eye movements from the subject (Staub, 2010) or longer fixation durations during regressive re-reading (P. C. Gordon, Hendrick, Johnson, & Lee, 2006).

Second, if a language has globally ambiguous RCs, comprehenders offer fewer ORC-interpretations relative to SRC-interpretations. Provided in (3) and (4) are examples of ambiguous relative clauses in Dutch (L. Frazier, 1987) and in Santiago Laxopa Zapotec (S. Foley et al., 2019). Also provided are their proportions of responses by parse (i.e., whether it was interpreted as an SRC or an ORC). L. Frazier collected the Dutch speakers' preferences using a comprehension question after a self-paced reading trial, while S. Foley et al. collected the Laxopeños' preference using a picture-selection task.

(3) de Amerikaanse die de Nederlander wil uitnodigen
the American who the Dutch wants to invite
'the American that wants to invite the Dutch person' *SRC* (74%)

'the American that the Dutch person wants to invite' *ORC* (26%)

(4) bez=e'nh tsyi'in beku'=nh xhan yage'=nh
fox=DEF bite.CONT dog=DEF under tree=DEF
'the fox that is biting the dog under the tree' *SRC* (62%)

'the fox that the dog is biting under the tree' *ORC* (38%)

This asymmetry seems to hold quite robustly across typologically different languages using a variety of behavioral and neural measures, in both child- and adult-langauges: Avar: (Polinsky, Gallo, Graff, & Kravtchenko, 2012); Chamorro: (Wagers et al., 2018); Chinese: (Vasishth, Chen, Li, & Guo, 2013; F. Wu, Kaiser, & Andersen, 2012); Dutch (L. Frazier, 1987; Mak, Vonk, & Schriefers, 2002, 2006); Ch'ol and Q'anjob'al: (Clemens et al., 2015); English (Gibson et al., 2005; Grodner & Gibson, 2005; King & Just, 1991; King & Kutas, 1995; Traxler, Morris, & Seely, 2002; Traxler, Williams, Blozis, & Morris, 2005); French (Cohen & Mehler, 1996); German (Bader & Meng, 1999); Georgian (S. Foley & Wagers, 2017); Hebrew (Arnon, 2010); Italian (Arosio, Guasti, & Stucchi, 2011); Japanese (Ueno & Garnsey, 2008); Korean (Kwon et al., 2010); Russian (Levy, Fedorenko, & Gibson, 2013); Spanish (Betancort, Carreiras, & Sturt, 2009); Tagalog (Pizarro-Guevara, 2014; Bondoc et al., 2018; Tanaka et al., 2019).

5.2 Accounting for the asymmetry

There are a number of proposals in the literature that have been argued to account for this well-known asymmetry. Broadly speaking, these proposals can be grouped into two classes, depending on what aspect of the dependency they emphasize. Some proposals emphasize the role of intervening linguistic material. Others emphasize the role of

linguistic experience.

This section is organized as follows. I first discuss the role of intervening material, which subsumes both structure- and memory-based accounts. I then discuss the role of linguistic experience, which subsumes accounts that emphasize how similar relative clauses are to main clauses, and accounts that highlight the importance of the frequency.

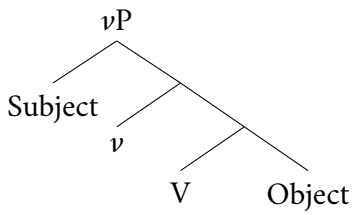
5.2.1 The role of intervening material

There are two flavors of intervention-based proposals: (i) those that emphasize structural locality; and (ii) those that emphasize the role of working memory, whose constructs are operationalized in terms of linear locality. While these proposals vary in terms of how they define what is considered intervening linguistic material, the basic intuition behind how processing difficulty is quantified remains the same. Intervention-based proposals relate processing difficulty to the number of intervening elements, such that the greater the number of intervening elements, the more difficult the dependency is to process.

Structural locality

Proposals that emphasize structural locality define intervention in terms of the number of syntactic projections intervening between the head noun and the gap (Hawkins, 1999; O'Grady, Lee, & Choo, 2003). Under this view, the subject-object asymmetry obtains because of how we combine the verb and its co-arguments. By hypothesis, subjects are generated in a structurally higher position than objects, as schematized in (5).

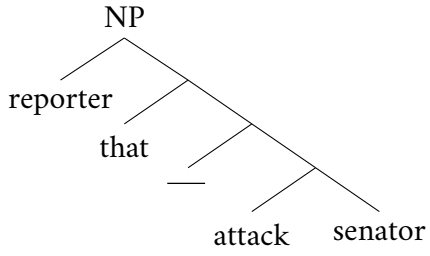
(5)



Different proposals vary in their assumptions about how articulated the clausal spine is. However, in a sense, how articulated it is has very little effect on how structural locality is quantified. The exact number of intervening projections will vary from proposal to proposal, but the overall pattern remains consistent: the number of intervening projections that ORCs have will always be greater than that of SRCs. This difference is schematized below: the distance between *the reporter* (the head) and the gap ("___") is greater in (7) compared to in (6).

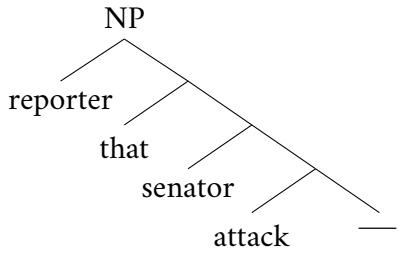
(6) a. The reporter that ___ attacked the senator … *SRC*

b.



(7) a. The reporter that the senator attacked ___ … *ORC*

b.



Assuming that how we combine verbs and their co-arguments is a universal feature of language, then these classes of proposal predict an asymmetry in any language.

Linear locality

These proposals define intervention in terms of the linear distance between the head noun and the gap (P. C. Gordon, Hendrick, & Johnson, 2001; Hsiao & Gibson, 2003; Grodner & Gibson, 2005; Lewis & Vasishth, 2005; Van Dyke & McElree, 2006; Carreiras, Duñabeitia, Vergara, de la Cruz-Pavía, & Laka, 2010). These proposals frame the effects

of linear distance as a function of working memory capacity. The main intuition behind these proposals is that increasing the linear distance imposes a greater burden on our working memory capacity, which then leads to greater processing difficulty. Under this view, the subject-object asymmetry obtains because there are more elements that linearly intervene between the head noun and the gap in ORCs than in SRCs.

There are various ways researchers have framed how linear distance affects working memory capacity and thus, how the subject-object asymmetry obtains. First, linear distance can affect storage costs. For example, in Dependency Locality Theory (DLT; Gibson, 1998), comprehenders keep track of incomplete/open dependencies in real-time and each incomplete/open dependency is associated with a storage cost. Under this view, ORCs are more costly in terms of memory storage because there are more incomplete/open dependencies to keep track of. Let's consider the example in (2) again, repeated here as (8).

(8) Adapted from Kwon et al. (2010)

a. The reporter that ___ attacked the senator disliked the editor *SRC*

b. The reporter that the senator attacked ___ disliked the editor *ORC*

In (8-a), when comprehenders have encountered *the reporter that attacked*, at this point in time, there are two incomplete dependencies: (i) the verb *attack* inside the relative clause requires an object DP; and (ii) the subject DP in the matrix clause, *the reporter*, requires a verb in the matrix clause. Meanwhile, in (8-b), when comprehenders have encountered *the reporter that the*, at this point in time, there are three incomplete dependencies—even though comprehenders are considering the same number of words as last time: (i) the determiner *the* inside the relative clause requires a noun; (ii) there is also a need for a verb inside the relative clause; and (iii) much like (2-a), the subject DP in the matrix clause requires a verb in the matrix clause.

Second, linear distance can affect integration cost. That is, it can affect how easily

we can establish the connection between an incoming word and the current mental representation of what is being processed. It is easier to establish the connection between elements in close (temporal) proximity than it is to establish the connection of elements in far (temporal) proximity. Under this view, SRCs are less costly in terms of integration than ORCs because the head noun and the gap are in closer proximity in SRCs than in ORCs.

Lastly, linear distance can also affect the retrieval process indirectly. Language comprehension requires accessing representations of items in memory that may not be what is currently being processed. This is especially true in non-local dependencies, like RCs. In a cue-based architecture of memory, the retrieval process is susceptible to SIMILARITY-BASED INTERFERENCE (P. C. Gordon et al., 2001; Van Dyke & Lewis, 2003; Lewis & Vasishth, 2005; Van Dyke & McElree, 2006, *inter alia*). Under this view, processing difficulty is modulated by how similar the items are in memory, and not by linear distance per se. Increasing the linear distance between the head noun and the gap can increase processing difficulty because it could add more items in memory that could compete with the head noun during retrieval. Framed this way, ORCs are more difficult to process than SRCs because in ORCs, the subject DP competes with the object head noun during the retrieval process because they have some overlapping features (e.g., they are both [+ANIMATE] referential expressions).

One key prediction of proposals that emphasize the role of similarity-based interference in the subject-object asymmetry is that this asymmetry will be reduced or leveled if the items in memory become more dissimilar. There is evidence suggesting that the asymmetry is indeed reduced when the intervening element is dissimilar. For example, P. C. Gordon, Hendrick, and Johnson (2004, *inter alia*) compared SRCs and ORCs by varying the referential type of the co-argument, as in (16). They found when the intervening co-argument in ORCs is the proper name *Bob*, the pronoun *you*, or the quantified expres-

sion *everyone*, the asymmetry is reduced in both reading times and accuracy.

(9) Adapted from P. C. Gordon and Lowder (2012, p. 406)

a. The senator that bothered (the reporter/Bob/you/everyone) caused … *SRC*

b. The senator that (the reporter/Bob/you/everyone) bothered caused … *ORC*

5.2.2 The role of linguistic experience

The main premise of these proposals is that the way in which comprehenders parse sentences is shaped by their prior exposure to their language. One class of proposals emphasizes the role of how frequent RCs are in the language, which in turn influences how comprehenders generate expectations in terms of what are likely continuations of a sentence. Another class of proposals emphasizes how similar RCs are to main clauses, which highlights the comprehenders' experience with how one would typically encounter the co-arguments in a clause.

Frequency

Frequency-based proposals argue that the ease of processing RCs is correlated with the relative abundance of the type of RCs in the language (Mitchell, Cuetos, Corley, & Brysbaert, 1995; Brysbaert & Mitchell, 1996; Reali & Christiansen, 2007). Under this view, ORCs are more difficult to process than SRCs because ORCs are not as frequent in the language.

Let's consider the subject-object asymmetry in English. Frequency-based explanations are tenable in English because ORCs are less frequent compared to SRCs in the language. Roland, Dick, and Elman (2007) analyzed 5 English corpora—the British National Corpus (BNC), BNC-Spoken, Brown, Switchboard, and Wall Street Journal Treebank—and found that ORCs were indeed less frequent compared to SRCs across the 5 corpora.

I provide a summary of the relevant results in Table 5.1.

Table 5.1: Frequencies of each type of relative clause in English, per 1 million noun phrases, as reported by Roland, Dick, and Elman (2007)

Type	BNC	BNC-Spoken	Brown	Switchboard	Wall Street Journal
SRC	14,182	9,851	15,024	9,548	18,229
ORC	2,943	3,863	1,976	5,616	1,802

Frequency-linked difficulties have been formalized in the literature in terms of surprisal (Levy, 2006). As comprehenders, we generate expectations in terms of what continuations are possible. Framed this way, processing difficulty is inversely related to the degree to which the input is consistent with our expectations. More concretely, after encountering *that* in (10), an SRC-continuation, as in (10-a), is much more likely given that SRCs are more frequent in the language. Thus, SRCs are preferentially expected by comprehenders. On the other hand, an ORC-continuation, as in (10-b), is more surprising given that ORCs are not as frequent in the language, and thus, is more difficult to process.

(10) The reporter that ...

a. attacked the senator disliked the editor *SRC-continuation = low surprisal*

b. the senator attacked disliked the editor *ORC-continuation = high surprisal*

Main clause similarity

These classes of proposals relate processing difficulty as a function of how similar the order of the RC is to the order of the main clause (Bever, 1970; Diessel & Tomasello, 2005). The more similar they are, the easier it is to process. Let's consider English once again. Shown in (11) is the order of the main clause in English: S-V-O. Notice that the order of elements in an SRC closely resembles this order, as shown in (12-a). The order of elements in an ORC, however, does not resemble it, as shown in (12-b). Under this view,

the asymmetry obtains in English because the order of elements in SRCs (i.e., Head-V-O) closely resembles the order of the main clause in the language (i.e., S-V-O).

(11) The reporter attacked the senator. *Main clause order: S-V-O*

(12) a. The reporter that attacked the senator … *RC order: Head-V-O*

b. The reporter that the senator attacked … *RC order: Head-S-V*

5.3 Attenuating the asymmetry

A related strand of research examines the various factors that can modulate this asymmetry. Here I survey three factors that have been found to reduce the asymmetry: (i) the animacy of the head noun; (ii) the referential type of the intervening element; and (iii) the relative order of the head noun and the relative clause.

5.3.1 Animacy of the head noun

Researchers have found that when the animacy of the head noun varies, such that ORCs have inanimate head nouns, the asymmetry is reduced or neutralized. For example, in Dutch, Mak et al. (2002) found that an ORC with an inanimate head noun, as in (13-b), did not show longer reading times compared to an SRC with an animate head noun, as in (13-a).

(13) Adapted from Mak et al. (2006)

a. de wandelaars die de rots beklommen hebben
the hikers that the rocks climbed have
'The hikers that climbed the rocks' *SRC*

b. de rots die de wandelaars beklommen hebben
the rocks that the hikers climbed have
'The rocks that the hikers climbed' *ORC*

F. Wu et al. (2012) found comparable results in Chinese. An ORC with an inanimate head noun, as in (14-b), was equally as easy to process as an SRC with an animate head noun, as in (14-a).

(14) Adapted from F. Wu et al. (2012)

a. duokai shikuai de jizhe chenggongde hunrule yingdi
dodge stone DE reporter successfully slipped into camp
'The reporter that dodged the stone successfully slipped into …' *SRC*

b. jizhe duokai de shikuaii zhongzhongde luozaile dishang
reporter dodge DE stone heavily fell to ground
'The stone that the reporter dodged fell to the ground heavily.' *ORC*

Comparable results were also found in English (Traxler et al., 2005; Gennari & MacDonald, 2008; Lowder & Gordon, 2014, *inter alia*). For example, Lowder and Gordon (2014) replicated the well known subject-object asymmetry with relative clauses where both the head noun and the co-argument were animate: (15-b) was harder to process than (15-a). They also replicated the effect of animacy on the reduction of the asymmetry: (15-b) was harder to process than (15-c).

(15) Adapted from Lowder and Gordon (2014)

a. The senator that criticized the journalist accused the governor … *SRC*

b. The senator that the journalist criticized accused … *ORC-Animate*

c. The article that the journalist criticized accused … *ORC-Inanimate*

5.3.2 Referential type of intervening elements

Some memory-based accounts, as discussed in Section 5.2.2, have examined the effects of the similarity of the items on retrieval. As items in memory become more dissimilar, the easier it would be to retrieve the representation of the intended item. One

key prediction of proposals that emphasize the role similarity-based interference in the subject-object asymmetry is that this asymmetry will be reduced or leveled if the items in memory become more dissimilar. Dissimilar items in memory would allow comprehenders to correctly retrieve the representation of the head noun, which in turn, could reduce the subject-object asymmetry. There is evidence that varying the referential type of the intervening elements facilitates the retrieval of the head noun in ORCs in English (P. C. Gordon et al., 2004, *inter alia*). P. C. Gordon et al. (2004) compared SRCs and ORCs by varying the referential type of the co-argument, as in (16). They found that when the intervening co-argument in ORCs is the proper name *Bob*, the pronoun *you*, or the quantified expression *everyone*, the asymmetry is reduced in both reading times and accuracy.

(16) Adapted from P. C. Gordon and Lowder (2012, p. 406)

a. The senator that bothered (the reporter/Bob/you/everyone) caused … *SRC*

b. The senator that (the reporter/Bob/you/everyone) bothered caused … *ORC*

The referential types of the head noun and the co-argument can also reduce the asymmetry in Russian. Price and Witzel (2017) compared SRCs and ORCs with full noun phrase co-arguments, as in (17-a) and (17-b), respectively, in experiment 1. They then compared SRCs and ORCs with pronominal co-arguments, as in (17-c) and (17-d), respectively, in experiment 2. In both experiments, they did not find the asymmetry in the participants' reading times. However, they did find the asymmetry in the participants' accuracy when the co-arguments were full noun phrases but not when they were pronouns.

(17) Adapted from Price and Witzel (2017)

a. Hozjajka, kotoraja posle progulki starušku sil'no rasstroila
Housewife who.NOM after walk old lady.ACC really upset
novostjami, legla …
with news lay …

'The housewife who after the walk really upset the old lady with the news lay (on the couch in the living room.)" *SRC: full NP co-argument*

b. Hozjajka, kotoraju posle progulki staruška sil'no rasstroila
Housewife who.NOM after walk old lady.NOM really upset
novostjami, legla …
with news lay …
'The housewife who after the walk the old lady really upset with the news lay (on the couch in the living room.)" *ORC: full NP co-argument*

c. Hozjajka, kotoraja posle progulki nas sil'no rasstroila novostjami,
Housewife who.NOM after walk 1PL.ACC really upset with news
legla …
lay …
'The housewife who after the walk really upset us with the news lay (on the couch in the living room.)" *SRC: pronominal co-argument*

d. Hozjajka, kotoraju posle progulki staruška sil'no rasstroila novostjami,
Housewife who.NOM after walk 1PL.NOM really upset with news
legla …
lay …
'The housewife who after the walk we really upset with the news lay (on the couch in the living room.)" *ORC: pronominal co-argument*

5.3.3 Relative order of the head noun and relative clause

While the subject-object asymmetry enjoys cross-linguistic and cross-methodological robustness, a high proportion of the data come from languages with head-initial RCs (Kwon et al., 2010). The studies on languages that have head-final RCs have had mixed results. Some show a clear asymmetry in the expected direction. Some, in the other direction. Others find no clear asymmetry. Thus, when we compare languages that have head-initial RCs and languages that have head-final RCs, there is a tenuous generalization that can be made. The relative order of the head noun and the RC seems to modulate the asymmetry. Here I survey languages with head-final RCs that do not seem to (always)

exhibit the expected asymmetry.

First, Chinese languages are notorious in the RC processing literature for having mixed results. In one of the earliest studies on Chinese RCs, Hsiao and Gibson (2003) conducted a self-paced reading study to compare the processing of SRCs, as in (18-a), and ORCs, in (18-b). Their results showed an asymmetry, albeit in the other direction: ORCs were easier to process than SRCs.

(18) Adaped from Hsiao and Gibson (2003)

a. [yaoching fuhao] de guanyuan shinhuaibugui danshi shanyu
invite tycoon DE official have bad intentions but good at
yintsang
hide
'The official that invited the tycoon has bad intentions but good at hiding them.' *SRC*

b. [fuhao yaoching] de guanyuan shinhuaibugui danshi shanyu
tycoon invite DE official have bad intentions but good at
yintsang
hide
'The official that the tycoon invited has bad intentions but good at hiding them.' *ORC*

Other self-paced reading studies also found that ORCs were easier to process than SRCs (B. Chen, Ning, Bi, & Dunlap, 2008; Lin & Garnsey, 2010). This was true even when the RCs were preceded by biasing context (Gibson & Wu, 2013). This finding was replicated in an ERP study (Packard, Ye, & Zhou, 2010), a study using the maze task (Qiao, Shen, & Forster, 2011), and an eye-tracking study (Sung, Tu, Cha, & Wu, 2016).

However, there are studies that found the expected asymmetry. In 2 of the 3 self-paced reading experiments conducted by Vasishth et al. (2013), SRCs were easier to process than ORCs. They also conducted a Bayesian random effects meta-analysis of 15 studies (including their 3 experiments) and maintained that the evidence weighs in favor of

SRCs being easier to process than ORCs. They argued that the variability in the studies is due in part to temporal ambiguity. In their self-paced reading and eye-tracking studies, Jäger, Chen, Li, Lin, and Vasishth (2015) also found that SRCs were easier to process than ORCs. Crucially, their experimental items apparently eliminated the temporal ambiguity that plagued previous studies. For a thorough review of the previous studies comparing SRCs and ORCs in Chinese, see Bulut, Cheng, Xu, Hung, and Wu (2018).

Second, in Basque, there is evidence for a reversal of the asymmetry. Carreiras et al. (2010) conducted two self-paced reading studies and one ERP experiment to compare SRCs with an ergative gap (i.e., SRCs with transitive verbs), as in (19-a), and ORCs with an absolutive gap, as in (19-b). They leveraged the homophony of a definite noun's absolutive plural (*–ak*) and its ergative singular form (*–a–k*) to create string-identical items that are only disambiguated upon encountering the very last word.

(19) Adapted from Carreiras et al. (2010)

a. [Irakasle-ak aipatu ditu-en] ikasle-a-k lagun-ak
teacher-PL mentioned have:3SG>3PL-REL student-DET-ERG friend-PL
ditu
have:3SG>3PL
'The student that mentioned the teacher has friends' *SRC*

b. [Irakasle-a-k aipatu ditu-en] ikasle-ak lagun-ak
teacher-DET-ERG mentioned have:3SG>3PL-REL student-PL friend-PL
dira
be:3PL
'The students that the teacher mentioned are friends' *ORC*

They found significantly slower reading times in the SRC condition than in the ORC condition at the disambiguating region (i.e., at the final word). Their EEG study further corroborated their findings: the SRC condition had a significantly larger P600 than the ORC condition. This measure is often associated with syntactic processing difficulty.

Third, in Avar, a Caucasian language spoken in Dagestan, Russia, Polinsky et al. (2012)

used self-paced reading to compare SRCs with an ergative gap (i.e., SRCs with transitive verbs), as in (20-a), and ORCs with an absolutive gap, as in (20-b). They found (i) a slow down upon encountering the first noun in the SRC condition; and (ii) no slow-down at the region after the head noun. If we are being generous, we could interpret these results as an initial reversal of the asymmetry early on, followed by neutralization later on. If we are being conservative, there is evidence for neutralization—at the very least.

(20) Adapted from Polinsky et al. (2012)

a. [ʕoloqana-y yas repetici-yal-de y-ač:-un y-ač'-ara-y]
unmarried-II girl.ABS rehearsal-OBL-LOC II-bring-GER II-come-PRTCP-II
artitska bercina-y y-igo
artist.ABS beautiful-II II-AUX
'The actress that brought the young girl to the rehearsal is pretty.' *SRC*

b. [xalq'iya-y artistka-yał repetici-yal-de y-ač:-un
distinguished-II actress-ERG rehearsal-OBL-LOC II-bring-GER
y-ač'-ara-y] yas bercina-y y-igo
II-come-PRTCP-II girl.ABS beautiful-II II-AUX
'The girl the distinguished actress brought to the rehearsal is pretty.' *ORC*

Some languages show no clear asymmetry, like Chinese. Other languages show some sort of neutralization, like Avar. Other languages show a reversal, like Basque. These results are consistent with the tenuous generalization formulated above: The relative order of the head noun and the RC seems to modulate the asymmetry.

It is important to keep in mind, however, that this generalization was formulated by comparing *different* languages. A skeptic might argue that the purported effect of word order on the modulation of the asymmetry is confounded by the fact that we are comparing different languages. In other words, it is not easy to divorce the independent contribution of word order on the modulation of the asymmetry from language-specific properties. A stronger case could be made for word order effects if language is held constant. What is needed then is a language where it is possible to have both head-initial and head-final

RCs—a typologically rare property for a language to have (Dryer, 2013). Any observed differences can then be attributed to "purer" word order effects.

More recently, Wagers et al. (2018) compared how SRCs and ORCs are processed in Chamorro, an Austronesian language of the Marianas. One crucial feature of Chamorro grammar is that it allows for both head-initial and head-final RCs, as in (21-b) and (21-a), respectively.

(21) Adapted from Wagers et al. (2018, p. 209)

a. Kao siña un li'i' atyu na tåotao [i matåta'chung]?
Q can AGR see that LNK person COMP AGR.sit.PROG
'Can you see the man who is sitting down?' *Head-initial*

b. Kao siña un li'i' atyu i [matåta'chung] na tåotao?
Q can AGR see that the AGR.sit.PROG LNK person
'Can you see the man who is sitting down?' *Head-final*

This feature allowed Wagers et al. (2018) to investigate how RCs were processed in both word orders *language internally*. They found the following. First, in head-initial RCs, Chamorro comprehenders preferred SRCs over ORCs. This echoes the well-known subject-object asymmetry that has been found to be cross-linguistically robust. Second, in head-final RCs, Chamorro comprehenders displayed a modest preference for ORCs over SRCs. This echoes the effect of word order on the modulation of the asymmetry. Finally, even though comprehenders ultimately preferred ORCs in head-final configurations, they were still faster at providing the correct interpretation when it was an SRC. What we can learn from this language internal comparison is that even when word order effects modulate the asymmetry, it nevertheless leaves a lingering "footprint" in our processing behavior.

5.4 The present studies: An overview

I conducted three experiments at the University of the Philippines – Diliman in the summer of 2019. These were run simultaneously in one session. The first experiment, Experiment 6, compared how unambiguous SRCs and ORCs were processed in both head-initial and head-final configurations. The 16 experimental items involved referential expressions as co-arguments. This design maximizes the comparability of my findings to other studies since most studies investigating how SRCs and ORCs are processed also use co-arguments that are referential expressions. Perhaps more importantly, this maximizes the comparability of my findings and the findings of Wagers et al. (2018), who, to my knowledge, were the first to investigate the effect of word order on the RC-processing asymmetry language internally.

A skeptic of whatever the results of Experiment 6 may be might argue that the naturalness of head-final RCs (or lack thereof) could be a confound. More specifically, as discussed in Chapter 2.7.2, head-final RCs—especially with full noun phrase co-arguments—are more marked and are not as natural sounding compared to head-initial RCs. In order to preemptively address such concerns, I had made the RCs "smaller" to make them as natural sounding as possible. The second experiment, Experiment 7, was essentially the same as the first, with one crucial difference. The 16 experimental items involved pronominal co-arguments. Not only does this design address the potential concern, it also allows us to investigate the potential modulatory effects of having co-arguments and head nouns be of different referential types à la P. C. Gordon et al. (2004).

The third experiment, Experiment 8, was essentially like the first, with two crucial differences. First, the 16 experimental items involved sentences with verbs in the recent perfective aspect. This grammatical aspect is standardly assumed to lack voice morphology, and both co-arguments of a semantically transitive verb are *typically* assigned genitive case marking (Kroeger, 1993, *inter alia*): Verb[RP]–*ng* NP–*ng* NP. These prop-

erties make sentences involving recent perfectives globally ambiguous. Second, I leveraged these quirky properties to compare how ambiguous and unambiguous RCs were processed. As we will see below, both unambiguous and ambiguous relative clauses are string-identical until the determiner of the co-argument, which is the point of disambiguation. This design allows for the second flavor of the asymmetry to be investigated, as we saw in Dutch (L. Frazier, 1987) and Santiago Laxopa Zapotec (S. Foley et al., 2019): if a language has globally ambiguous RCs, comprehenders offer fewer ORC-interpretations relative to SRC-interpretations.

These three experiments were interspersed with a set of 24 items looking at relative clauses that use a different relativization strategy as the ones above. These were included primarily to add more variety to the types of relative clauses that participants saw throughout the experiment. They will not be analyzed in this dissertation.

5.4.1 Participants

I recruited 65 speakers from UP – Diliman, using the same recruitment criteria discussed in the previous experiments. They ranged from 18 to 59 years of age (M = 25, SD = 8), and all lived in and around Metro Manila at the time of testing. They received either a gift certificate from Bo's (a local coffee shop) or Fully Booked (a local bookstore) valued at 300 Philippine Pesos, for participating in the study.

5.4.2 Procedure

The experiment was deployed in OpenSesame (Mathôt, Schreij, & Theeuwes, 2012), using a Surface Pro tablet. A schematization of a typical trial in Experiments 6–9 is provided below in Figure 5.1. There are three main events for any given trial: (i) Context presentation; (ii) Picture selection; and (iii) Confidence rating. These are described in greater detail below.

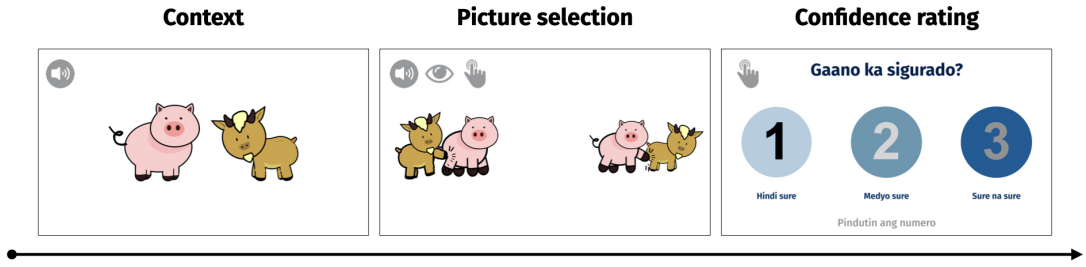


Figure 5.1: A schematization of typical trial in Experiments 6–8. There are three main events per trial: (i) Context presentation, where participants saw a context screen and heard a recording introducing the characters associated with this trial; (ii) Picture selection task, where participants saw the two characters perform an action to each other, and where they heard a recording containing an RC and chose the picture best representing the recording they had just heard; and (iii) Confidence rating, where participants had to rate how confident they were with their selection.

First, participants saw a context screen that contained two animals (Experiments 6, and 8) or one animal (Experiment 7). They then heard a recording of a context that introduced the character(s) associated with this particular trial. This context was also a way to get the participants acquainted with the verb used in this trial. Crucially, the way in which verbs were inflected in the context screen never matched how participants had heard them embedded inside an RC. The context script used in Experiments 6 and 8 is provided in (22) and that used in Experiment 7 is provided in (23)

(22) May isa-ng [ANIMAL 1] at isa-ng [ANIMAL 2]. Minsan gusto
EXIST one-LNK and one-LNK sometimes want
nila-ng maN[VERB]. Minsan naman, gusto nila-ng
3PL.GEN-LNK sometimes on the other hand want 3PL.GEN-LNK
magpa[VERB]

There is a [ANIMAL 1] and a [ANIMAL 2]. Sometimes, they like to [VERB]. Sometimes, they liked to be [VERB]ed. *Exps. 6, 8*

(23) May isa-ng [ANIMAL 1]. Minsan gusto niya-ng maN[VERB].
EXIST one-LNK sometimes want 3SG.GEN-LNK
Minsan naman, gusto niya-ng magpa[VERB]
sometimes on the other hand want 3SG.GEN-LNK

There is a [ANIMAL 1]. Sometimes, s/he likes to [VERB]. Sometimes, s/he likes to be [VERB]ed. *Exp. 7*

Second, after the offset of the context recording, participants saw another screen that contained two pictures in which two animals are performing the same action to each other. They then heard a recording containing the experimental item. They were to choose the picture that best represented what they had just heard. They had three seconds to choose a picture after the offset of the recording; otherwise, it would time out. The experimental items will be provided in their respective sections.

Finally, after they had chosen a picture (or not if they timed out), participants had to rate how confident they were with their selection. They were explicitly instructed that if they saw this screen even if they did not choose a picture (i.e., they timed out), they were to press 1 *hindi sure* "not sure."

5.4.3 Data analysis

I used two measures to compare how SRCs and ORCs were processed. The first measure was the participants' choice data. The second measure was how quickly participants offered a correct interpretation, their RT data.

To get a more nuanced picture of their choice data, I took into account what they had chosen (i.e., whether they chose the correct or incorrect response for unambiguous RCs, and whether they chose an SRC-interpretation or an ORC-interpretation for ambiguous RCs) and their confidence rating. I derived an ordinal scale with 6 possible responses. For ambiguous RCs, the six derived responses possible were as follows: "Incorrect, really confident," "Incorrect, somewhat confident," "Incorrect, not confident," "Correct, not confident," "Correct, somewhat confident," and "Correct, confident." For ambiguous RCs, since by definition, there cannot be a correct response, the six derived answers possible were as follows: "ORC, really confident," "ORC, somewhat confident," "ORC, not confind-

ent," "SRC, not confident," SRC somewhat confident," and "SRC, really confident."

Before proceeding, I make explicit my exclusion criteria and the percentage of data that were lost from exclusion. I excluded trials where participants had timed out (i.e., 3 seconds after audio offset of the picture-selection event) and where they clicked outside the regions of interest. Provided in Table 5.2 is a summary of the excluded data.

Table 5.2: Summary of excluded data (in %) by exclusion-type for Experiments 6–9

	Timeouts	Clicks outside ROI	Total
Experiment 6	3.67	.10	3.77
Experiment 7	.87	.10	.97
Experiment 8	4.27	.10	4.37

For the choice data, I estimated cumulative link mixed models using the *Ordinal* package (Christensen, 2018) in R (R Core Team, 2018). For the RTs, I estimated linear mixed effects models using *lme4* (Bates et al., 2015) in R (R Core Team, 2018). I then included the maximal random effects structure justified by the design that allowed the models to converge (Barr et al., 2011; Barr, 2013). I will include the exact specifications of the maximal models that converged when I report the results for each experiment.

5.4.4 Preview of the results

Below I preview the results of the upcoming experiments. While the results are quantitatively different, they are qualitatively consistent across 3 experiments: even when word order attenuated the asymmetry in the choice data, the asymmetry nevertheless persisted in the RT-data.

Table 5.3: Preview of the results by experiment and dependent measure. A "✓" represents consistency with the expected asymmetry. A "✗" represents inconsistency with the expected asymmetry, either via neutralization or reversal.

Experiment		Choice	RT
Exp. 6: Unambiguous RCs with R-expression co-arguments	Head-initial	✓	✓
	Head-final	✗	✓
Exp. 7: Unambiguous RCs with pronominal co-arguments	Head-initial	✗	✓
	Head-final	✗	✓
Exp. 8: Ambiguous RCs with R-expression co-arguments	Head-initial	✓	✓
	Head-final	✗	✗

5.5 Experiments 6: Comparing unambiguous RCs with R-expression co-arguments

This section compared how unambiguous SRCs and ORCs were processed in both head-initial and head-final configurations. The verbs exhibited either AV or PV and the co-arguments were referential expressions. In a nutshell, I observed an asymmetry in both the choice and RT data when the RC was head-initial. The asymmetry was neutralized in the choice data when the RC was head-final. It persisted, however, in the RT data.

5.5.1 Materials

The experiment employed a 2×2 design, crossing whether the relative clause was head-initial or head-final (ORDER: HI, HF) and whether the RC was an SRC or an ORC (PARSE: SRC, ORC). I created 16 items involving reversible predicates. Each item was distributed evenly across four lists via Latin Square design. The items of this experiment were combined together in one session and randomized with the 72 items from the other experiments. Provided in Table 5.4 is a sample experimental item.

Table 5.4: Sample experimental items used in Experiment 6, manipulating ORDER (HI, HF), and PARSE (SRC, ORC). The head noun is bolded and the RCs are demarcated by "[]."

ORDER	PARSE					
		May	isang	baboy	at	kambing
		EXIST	one.LNK	pig	and	goat
		'There is a pig and a goat.'				
		Minsan	gusto	nilang	manipa	
		sometimes	want	3PL.GEN.LNK	maN.kick	
		'Sometimes they like to kick.'				
		Minsan	naman	gusto	nilang	magpasipa
		sometimes	also	want	3PL.GEN.LNK	magpa.kick
		'Sometimes they also likes to be kicked.'				
		Piliin	ang	larawan	ng …	
		choose.PV	NOM	picture	GEN	
		'Choose the picture of …'				
HI	SRC	**baboy**	na	[sumisipa	ng	kambing]
		pig	LNK	kick.AV	GEN	goat
		'…the pig that is kicking the goat.'				
HI	ORC	**baboy**	na	[sinisipa	ng	kambing]
		pig	LNK	kick.PV	GEN	goat
		'…the pig that the goat is kicking.'				
HF	SRC	[sumisipa	ng	kambing]	na	**baboy**
		kick.AV	GEN	goat	LNK	pig
		'…the pig that is kicking the goat.'				
HF	ORC	[sinisipa	ng	kambing]	na	**baboy**
		kick.PV	GEN	goat	LNK	pig
		'…the pig that the goat is kicking.'				

5.5.2 Predictions

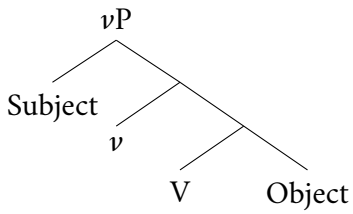
Structural locality

Recall that the structural locality view relates processing difficulty to the number of syntactic projections intervening between the head noun and the gap. The greater the number, the more difficult it is to process.

It is perhaps uncontroversial to assume that the mechanism by which we combine

the verb and its co-arguments is the same in English and Tagalog. That is, subjects are generated in a structurally higher position than objects, as schematized in (24).

(24)



Let's assume that the algorithm for counting intervening syntactic projections operates at the level where co-arguments are base-generated. This predicts an asymmetry, where SRCs will be processed easier than ORCs, because the number of intervening projections that ORCs have will always be greater than that of SRCs—no matter how articulated the assumed clausal spine is.

Linear locality

Recall that the linear locality view relates processing difficulty to the number of elements linearly intervening between the head noun and the gap. The greater the number, the more difficult it is to process.

When the verb exhibits AV, the order of the elements can either be VSO or VOS. Meanwhile, when the verb exhibits PV, the order of the elements can only be VSO. I use these word order possibilities to approximate the number of linear interveners between the head noun and the gap, represented by circled numerals, for head-initial and head-final RCs in (25) and (26), respectively.

(25) Head-initial RCs

a. Head ① Verb.AV ___ Object *Linear distance = 1*

b. Head ① Verb.AV ② Object ___ *Linear distance = 2*

c. Head ① Verb.PV ② Subject ___ *Linear distance = 2*

(26) Head-final RCs

a. Verb.AV ___ ① Object Head *Linear distance = 1*

b. Verb.AV Object ___ Head *Linear distance = 0*

c. Verb.PV Subject ___ Head *Linear distance = 0*

In head-initial configurations, the linear distance between the head noun and the gap in SRCs can either be 1 or 2, as in (25-a) and (25-b), respectively, depending on which word order is used to calculate linear distance. The linear distance in ORCs is 2, as in (25-c). A conservative estimate of linear distance predicts neutralization because the distances in (25-b) and (25-c) are equal. On the other hand, a more liberal estimate predicts the expected asymmetry because the distance in (25-b) is smaller than that in (25-c).

In head-final configurations, the linear distance between the head noun and the gap in SRCs can either be 1 or 0, as in (26-a) and (26-b), respectively, depending on which word order is used to calculate the distance. The linear distance in ORCs is 0, as in (26-c). A conservative estimate of linear distance predicts a reversal of the asymmetry because the distance in (26-b) is greater than that in (26-c). On the other hand, a more liberal estimate predicts neutralization because the distances in (26-b) and (26-c) are equal.

Frequency

Recall that the frequency view relates processing difficulty to be function of the relative abundance of the type of RCs in the language. In Chapter 2, I talked about the results of three corpus studies in Tagalog. Nagaya (2019) found that ORCs are more common than SRCs in Tagalog. His findings echo the general prevalence of PV-marked verbs in the language, relative to AV-marked (Cooreman et al., 1984; Garcia et al., 2019).

The prediction of a pure frequency-based account is fairly straightforward. Because ORCs are more frequent than SRCs, then it is predicted that ORCs will be easier to process than SRCs. Thus, a reversal is expected in both head-initial and head-final configurations.

Main clause similarity

Recall that the main clause similarity view relates processing difficulty to be a function of how similar the order of the RC is to the order of the main clause. The more similar they are, the easier it is to process. In Tagalog, voice and word order interact, such that when the verb exhibits AV, the order of the elements can either be VSO or VOS. Meanwhile, when the verb exhibits PV, the order of the elements can only be VSO.

If comprehenders compute processing difficulty by paying attention to S, V, and O, then it is unclear what the prediction would be since there is no main clause word order where S or O precedes V. Instead I assume that comprehenders compute processing difficulty by paying attention to the relative order of S and O in main clauses.[1] Table 5.5 provides a comparison of the relative order of S and O in main clauses and RCs.

Table 5.5: Comparison of the relative order of S and O in main clauses and relative clauses in Tagalog. A "✓" indicates that the order of elements in main clauses and relative clauses match. A "✗" indicates the order of the elements in main clause and relative clauses do not match.

	Main	Head-initial		Head-final	
Voice	clause	SRC	ORC	SRC	ORC
AV	S - O	✓	N/A	✗	N/A
	O - S	✗	N/A	✓	N/A
PV	S - O	N/A	✗	N/A	✓

This weaker version of the main clause similarity account predicts an asymmetry in head-initial RCs and neutralization in head-final RCs, as sketched out in Table 5.5. In head-initial RCs, comprehenders have access to a main clause order where S precedes O, which is the order of elements in a head-initial SRC. However, they have no access to a main clause order where O precedes S, which is the order of elements in a head-initial ORC. In head-final RCs, comprehenders have access to a main clause order where

[1] This weaker version of the main clause similarity can still account for the asymmetry in English. In English main clauses, S precedes O. This order is mirrored in SRCs. In ORCs, O precedes S. Under a main clause similarity view, this deviation of the relative order of S and O in ORCs causes processing difficulty.

S follows O, which is the order of elements in a head-final SRC. Similarly, they also have access to a main clause order where O follows S, which is the order of elements in a head-final ORC.

5.5.3 Data analysis

For the choice data, I estimated cumulative link mixed models using the *Ordinal* package in R. I used their derived ordinal ratings as the dependent measure, and the fixed effects were Order (hi, hf), Parse (src, orc), and their interactions. Order and Parse were sum-coded, such that hi and src mapped to the negative coefficients. I then included the maximal random effects structure justified by the design that allowed the models to converge.

For the RT data, I estimated linear mixed effects models using *lme4* in R. I used their log-transformed RT as the dependent measure, and the fixed effects were again Order (hi, hf), Parse (src, orc), and their interactions. Order and Parse were again sum-coded, such that hi and src mapped to the negative coefficients. I then included the maximal random effects structure justified by the design that allowed the models to converge

5.5.4 Results and discussion

Choice data

In Figure 5.2, I report the breakdown of their responses by accuracy and confidence. In Table 5.6, I provide the summary of the estimated model.

I found a main effect of Order, which suggests that participants were more accurate and confident in head-initial RCs than in head-final RCs. I also found a main effect of Parse, which suggests that participants were more accurate and confident in SRCs than in ORCs. These effects were qualified by a significant Order×Parse interaction. Participants were more accurate and confident in SRCs than in ORCs when the RC was

Figure 5.2: Breakdown of participants' responses by accuracy and confidence in Experiment 6. Negative values on the x-axis indicate incorrect responses, while positive values indicate correct responses. High values (3) indicate high confidence in their response (irrespective of accuracy), while low values indicate low confidence in their response.

head-initial than when the RC was head-final. Pairwise comparison of SRCs and ORCs in head-initial RCs indicate that SRCs were rated more accurately and confidently than ORCs (p = .01). Pairwise comparison of SRCs and ORCs in head-final RCs indicate no difference between SRCs and ORCs (p = .43)

RT data

In Figure 5.3, I provide a plot of the RTs of the participants' correct responses. In Table 5.7, I provide the summary of the estimated model.

I found a main effect of Order, which suggests that participants were faster at offering correct interpretations when the RC was head-initial than when it was head-final. I also found a main effect of Parse, which suggests that participants were faster at offering correct SRC-interpretations than correct ORC-interpretations.

Table 5.6: Summary of cumulative link mixed models in Experiment 6. The models included derived choice ratings as the dependent measure, and sum-coded ORDER, PARSE, and their interactions into the models.

```
clmm(Rating~Parse * Order + (1 + Parse * Order | Subject ) + (1 + Parse + Order
| Item ))
```

	Estimate	*SE*	z
ORDER	-3.35	.66	-5.05**
PARSE	-1.72	.65	-2.65*
ORDER×PARSE	2.87	1.24	2.32*

*$p < .05$, **$p < .001$

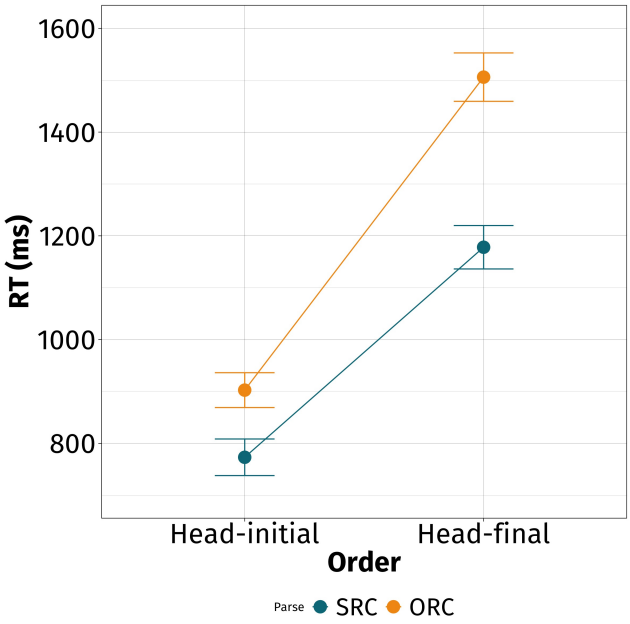


Figure 5.3: RTs of correct responses in Experiment 6.

Discussion

Analyses of the choice data suggest an asymmetry in head-initial RCs, but there is no evidence for an asymmetry in head-final RCs. Analyses of the RT data, on the other hand, suggest an asymmetry in both head-initial and head-final RCs.

How does each account fare? In Table 5.8, I provide an overview of how each account fared. A "✓" indicates that the prediction of a given account is consistent with the results,

Table 5.7: Summary of linear mixed-effects models in Experiment 6. The models included log-transformed RT as the dependent measure, and sum-coded ORDER, PARSE, and their interactions into the models.

```
lmer(log(RT)~Parse * Order + (1 + Parse + Order | Subject ) + (1 + Parse | Item
                                    ))
```

	Estimate	*SE*	z
Intercept	6.88	.06	112.07**
ORDER	.50	.05	10.87**
PARSE	.17	.06	2.76*
ORDER×PARSE	-.01	.07	-.15

*$p < .05$, **$p < .001$

while a "✗" indicates that it is not consistent. A "()" indicates that it may be consistent, but with some qualifications.

Overall, accounts emphasizing structural locality and main clause similarity fared quite well. These accounts are consistent with 75% of the results. Notice that for head-final RCs, where one account is inconsistent with the results, the other is consistent. If we take a more liberal estimate of linear distance, linear distance also fared quite well. Pure frequency-based approaches fared poorly, however. I return to this in the General Discussion.

One could argue that the predictions of frequency-based accounts are misrepresented here. For example, it could be that having referential expressions as co-arguments are not that frequent. There is evidence in English that in ORCs, it is much more frequent to have pronominalized co-arguments (Reali & Christiansen, 2007). I have no such data for Tagalog, but let's assume that the English facts also hold true for Tagalog. Another point of contention could be that having referential co-arguments is unnatural in Tagalog, as my field work indicated. Speakers seem to prefer smaller sized RCs in general. If we take this preference as a window into the intimate correlation between naturalness and frequency, then perhaps the experimental items used in Experiment 6 are not representative of what

Table 5.8: An evaluation of how each account fares compared against the results of Experiment 6. A "✓" indicates that the prediction of a given account is consistent with the results of Experiment 6. A "✗" indicates that it is not consistent with the results. A "()" indicates that it may be consistent, but with some qualifications.

Account	Order	Prediction	Choice	RT
Structural	Head-initial	Asymmetry	✓	✓
locality	Head-final	Asymmetry	✗	✓
Linear	Head-initial	Neutralization[a]/Asymmetry[b]	(✓)[b]	(✓)[b]
locality	Head-final	Reversal[a]/Neutralization[b]	(✓)[b]	✗
Frequency	Head-initial	Reversal	✗	✗
	Head-final	Reversal	✗	✗
Main clause	Head-initial	Asymmetry	✓	✓
similarity	Head-final	Neutralization	✓	✗

[a] is a more conservative estimate of linear distance; [b] is a more liberal estimate of linear distance

speakers actually hear/use. Thus, it is unwise to discount frequency-based approaches based on items that might not be ecologically valid. The next experiment, Experiment 7, attempts to address these potential concerns by having pronominalized arguments embedded in the RCs.

5.6 Experiments 7: Comparing unambiguous RCs with pronominal co-arguments

This section compared how unambiguous SRCs and ORCs were processed in both head-initial and head-final configurations. The verbs exhibited either AV or PV and the co-arguments were now pronominalized. In a nutshell, I observed no asymmetry in the choice data. The asymmetry, however, was present in the RT data.

5.6.1 Materials

The experiment employed a 2×2 design, crossing whether the relative clause was head-initial or head-final (ORDER: HI, HF) and whether the relative clause was an SRC or an ORC (PARSE: SRC, ORC). I created 16 items involving reversible predicates. Each item was distributed evenly across four lists via Latin Square design. The items of this experiment were combined together in one session and randomized with the 72 items from the other experiments. Provided in Table 5.9 is a sample experimental item.

5.6.2 Prediction

Structural locality

Let's continue to assume that the algorithm for counting intervening syntactic projections operates at the level where co-arguments are base-generated. This predicts an asymmetry, where SRCs will be processed easier than ORCs.

Linear locality

Recall that genitive-marked pronouns are second position clitics, as in (27).

(27) Kinakagat niya ang **buwaya**
bite.PV 3SG.GEN NOM crocodile
'S/he is biting the crocodile.'

Assuming that linear locality is computed using this word order, I approximate the number of linear interveners between the head noun and the gap, represented in circled numerals, for head-initial and head-final ORCs in (28).

(28) Linear distance in ORCs

a. Head ① Verb.PV Subj_{pro} ___ *Linear distance = 1*

Table 5.9: Sample experimental items used in Experiment 7, manipulating ORDER (HI, HF), and PARSE (SRC, ORC). The head noun is bolded and the RCs are demarcated by "[]."

ORDER	PARSE					
		May	isang	leon		
		EXIST	one.LNK	lion		
		'There is a lion.'				
		Minsan	gusto	niyang	mangagat	
		sometimes	want	3SG.GEN.LNK	maN.bite	
		'Sometimes s/he likes to bite.'				
		Minsan	naman	gusto	niyang	magpakagat
		sometimes	also	want	3SG.GEN.LNK	magpa.bite
		'Sometimes s/he also likes to be bitten.'				
		Piliin	ang	larawan	ng …	
		choose.PV	NOM	picture	GEN	
		'Choose the picture of …'				
HI	SRC	**buwaya**	na	[kumakagat	sa	kaniya]
		crocodile	LNK	bite.AV	DAT	3SG.DAT
		'…the crocodile that is biting him.'				
HI	ORC	**buwaya**	na	[kinakagat	niya]	
		crocodile	LNK	bite.PV	3SG.GEN	
		'…the crocodile that s/he is biting.'				
HF	SRC	[kumakagat	sa	kaniya]	na	**buwaya**
		bite.AV	DAT	3SG.DAT	LNK	crocodile
		'…the crocodile that is biting him.'				
HF	ORC	[kinakagat	niya]	na	**buwaya**	
		bite.PV	3SG.GEN	LNK	crocodile	
		'…the crocodile that s/he is biting.'				

b. Verb.PV Subj_{pro} ___ Head *Linear distance = 0*

Where it gets tricky is how to approximate the number of linear interveners between the head noun and the gap for SRCs primarily because "undoing" the movement in (29-a) yields an ungrammatical—or a very degraded—sentence, as in (29-b).[2] It is not about the word order since irrespective of where the argument *ang buwaya* surfaces, it is ungram-

[2]Thanks Henrison Hsieh, Kristina Gallego, Mantha Sadural, and Rowena Garcia for validating this contrast for me.

matical.[3] In addition, it is not directly about the pronominality of *kaniya* since a sentence involving an R-expression is also ungrammatical, as in (30-b). It is outside the scope of this dissertation to provide an analysis for why "undoing" this movement results to ungrammaticality. For the purposes of the dissertation, the empirical generalization will suffice: AV is incompatible with specific objects if the subject is not extracted, as in (29-b) and (30-b); AV can be compatible with them if the subject is extracted, as in (29-a) and (30-a). This contrast has previously been observed by Schachter and Otanes (1983), Maclachlan and Nakamura (1997), Aldridge (2004a), and others.

(29) a. Berde ang buwaya na kumagat sa kaniya
green NOM crocodile LNK bite.AV DAT 3SG.DAT
'The crocodile that bit him is green.'

b. */??Kumagat (ang buwaya) sa kaniya (ang buwaya)
bite.AV NOM crocodile DAT 3SG.DAT NOM crocodile
Intended: The crocodile bit him'

(30) a. Berde ang buwaya na kumagat sa bata
green NOM crocodile LNK bite.AV DAT child
'The crocodile that bit the child is green.'

b. */??Kumagat (ang buwaya) sa bata (ang buwaya)
bite.AV NOM crocodile DAT child NOM crocodile
Intended: The crocodile bit the child'

To be maximally explicit, I do not have a clear sense in terms of what accounts emphasizing linear locality will predict. If we are charitable and use the slight improvement in (29-b) when *ang buwaya* surfaces after the pronoun (see footnote 3) as the word order used to calculate linear locality, then some predictions can be made. They should, however, be taken with a grain of salt.

[3] I should note that there was a speaker who found it slightly better if *ang buwaya* is after *sa kaniya*. They then added that they would probably use PV instead.

(31) Linear distance in SRCs

a. Head ① Verb.AV sa Obj_{pro} ___ *Linear distance = 1*

b. Verb.AV sa Obj_{pro} ___ Head *Linear distance = 0*

Above I have approximated the number of linear interveners that an SRC can have—if we are being charitable. Assuming this to be true, then linear locality would predict neutralization in head-initial and head-final RCs because the distance between the head noun and gap in both RC-configurations is equal.

Alternatively, we can recast what linear distance is supposed to be a proxy of back into its original form, explanations that are memory-based. Even if we can't "undo" the movement in SRCs, we still have to retrieve the representation of the head noun in memory. Let's assume that retrieval is launched after we encounter the pronominal co-argument in both SRCs and ORCs. Under a similarity-based interference view, we do not expect the pronominal co-arguments to act as good interveners (P. C. Gordon et al., 2004). Not only are the head noun and co-argument different in referential type, they are also different in terms of case. Thus, we might expect neutralizations across the board.

Frequency

Recall that Nagaya (2019) found that ORCs are more common than SRCs in Tagalog, which echo the general prevalence of PV-marked verbs in the language, relative to AV-marked (Cooreman et al., 1984; Garcia et al., 2019). Also recall that in Experiment 1, I maintained that even though pure frequency-based accounts fared quite poorly, it is unwise to discount them just yet. One could argue that having referential expressions as co-arguments are uncommon in RCs and/or these "bigger" RCs are not frequent in Tagalog's input as "smaller" RCs.

While I have no frequency data in terms of what types of co-arguments are embedded inside, let's use known facts as starting points to speculate what the predictions would be.

If we are being conservative estimators, we can assume that the prevalence of verbs with PV inside RCs and in general reflects the frequency when the co-arguments are pronominalized, then we still expect reversals—much like in Experiment 6. If we are being liberal estimators, we can use Reali and Christiansen (2007) as a starting point. Because pronominalized co-arguments are more frequent in ORCs, we can assume that ORCs won't incur that great a processing cost now. By hypothesis, we might see neutralization now that ORC structures are more ecologically valid.

To reiterate, the predictions discussed here are speculative in nature: at best, we might expect neutralization; at worst, reversal.

Main clause similarity

Accounts emphasizing main clause similarity run into the same concerns as the accounts emphasizing linear distance. While the prediction for ORCs is relatively straightforward, fleshing out the prediction for SRCs is not as straightforward. The main reason is that "undoing" the movement in (32-a) yields an ungrammatical/extremely degraded sentence, as in (32-b). As mentioned in footnote 3, there was a speaker who found (32-b) better if *ang buwaya* surfaced after the pronoun.

(32) a. Berde ang buwaya na kumagat sa kaniya
green NOM crocodile LNK bite.AV DAT 3SG.DAT
'The crocodile that bit him is green.'

b. */??Kumagat (ang buwaya) sa kaniya (ang buwaya)
bite.AV NOM crocodile DAT 3SG.DAT NOM crocodile
Intended: The crocodile bit him'

To be maximally explicit, I do not have a clear sense in terms of what accounts emphasizing main clause similarity will predict. If we are charitable and take the slight improvement of (32-b) as the word order of the main clause, then there are predictions to be made.

These need to be taken with a grain of salt, however.

Assuming that we have the following word orders in the matrix clause, accounts emphasizing main clause similarity predict neutralization in head-initial RCs because there is no order where S precedes O when the verb exhibits AV, as in (33-a), and there is no order where O precedes S when the verb exhibits PV, as in (33-b). Similarly, these accounts would predict neutralization in head-final RCs because the head surfaces where S and O surface, respectively:

(33) a. Verb.AV sa Obj_{pro} Subj

b. Verb.PV Subj_{pro} Obj

5.6.3 Data analysis

For the choice data, I estimated cumulative link mixed models using the *Ordinal* package in R. I used their derived ordinal ratings as the dependent measure, and the fixed effects were ORDER (HI, HF), PARSE (SRC, ORC), and their interactions. ORDER and PARSE were sum-coded, such that HI and SRC mapped to the negative coefficients. I then included the maximal random effects structure justified by the design that allowed the models to converge.

For the RTs, I estimated linear mixed effects models using *lme4* in R. I used their log-transformed RT as the dependent measure, and the fixed effects were again ORDER (HI, HF), PARSE (SRC, ORC), and their interactions. ORDER and PARSE were again sum-coded, such that HI and SRC mapped to the negative coefficients. I then included the maximal random effects structure justified by the design that allowed the models to converge.

5.6.4 Results and discussion

Choice data

In Figure 5.6, I report the breakdown of their responses by accuracy and confidence. In Table 5.14, I provide the summary of the estimated model.

Figure 5.4: Breakdown of participants' responses by accuracy and confidence in Experiment 7. Negative values on the x-axis indicate incorrect responses, while positive values indicate correct responses. High values (3) indicate high confidence in their response (irrespective of accuracy), while low values indicate low confidence in their response.

There was no evidence that participants were more accurate and confident in head-initial RCs. There was also no evidence that participants were more accurate and confident in SRCs.

RT data

In Figure 5.5, I provide a plot of RTs of the participants' correct responses. In Table 5.11, I provide the summary of the estimated model.

I found a main effect of Parse, which suggests that participants were faster at offering correct SRC-interpretations than correct ORC-interpretations. This effect is qualified

Table 5.10: Summary of cumulative link mixed models in Experiment 7. The models included derived choice ratings as the dependent measure, and sum-coded ORDER, PARSE, and their interactions into the models.

```
clmm(Rating~Parse * Order + (1 + Parse * Order | Subject ) + (1 + Parse + Order
| Item ))
```

	Estimate	*SE*	z
ORDER	-.10	1.26	-.08
PARSE	-3.13	1.61	-1.94
ORDER×PARSE	.06	2.66	.02

$^{*}p < .05$, $^{**}p < .001$

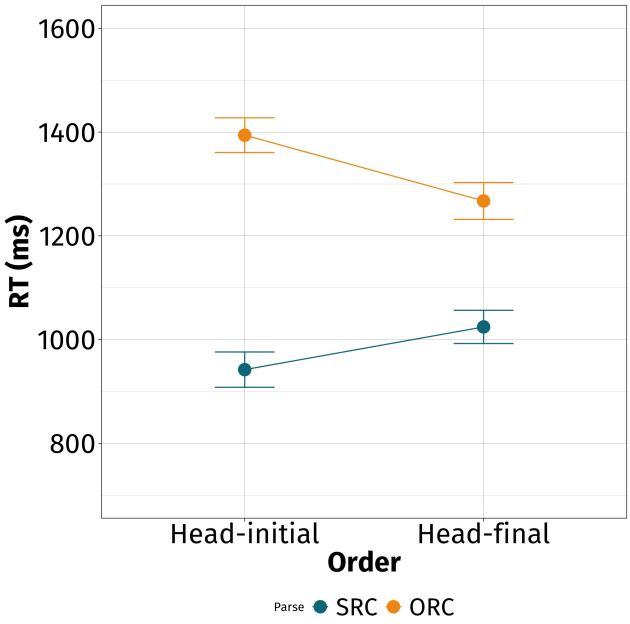


Figure 5.5: RTs of correct responses in Experiment 7.

by a significant ORDER×PARSE interaction. Participants were faster at offering correct SRC interpretations than ORC interpretations when the RC was head-initial than when it was head-final. Pairwise comparison of SRCs and ORCs in head-initial RCs indicate that they were offering correct SRC-interpretations faster than they were offering correct ORC-interpretations ($p < .001$). Pairwise comparison of SRCs and ORCs in head-final RCs indicate that they were also offering correct SRC-interpretations faster than they were

offering correct ORC-interpretations ($p < .001$).

Table 5.11: Summary of linear mixed-effects models in Experiment 7. The models included log-transformed RT as the dependent measure, and sum-coded ORDER, PARSE, and their interactions into the models.

```
lmer(log(RT)~Parse * Order + (1 + Parse | Subject ) + (1 + Parse | Item ))
```

	Estimate	*SE*	z
Intercept	6.92	.04	162.77**
ORDER	.004	.03	.13
PARSE	.38	.05	6.99**
ORDER×PARSE	-.27	.06	-4.92**

*$p < .05$, **$p < .001$

Discussion

Analyses of the choice data suggest no asymmetry in both head-initial and head-final RCs. Analyses of the RT data, on the other hand, suggest an asymmetry in both head-initial and head-final RCs.

How does each account fare? In Table 5.12, I provide an overview of how each account fared. A "✓" indicates that the prediction of a given account is consistent with the results, while a "✗" indicates that it is not consistent. A "()" indicates that it may be consistent, but with some qualifications.

Overall, accounts emphasizing structural locality fared relatively well. The neutralization in the participants' choice data is inconsistent with the predictions of structure-based accounts, but the lingering asymmetry in the participants' RT data is consistent. Accounts that have a mechanism for similarity-based interference also fared relatively well. Assuming that the findings in English ORCs extend to Tagalog as well (Reali & Christiansen, 2007), frequency-based explanations also fared relatively well.

Table 5.12: An evaluation of how each account fares compared against the results of Experiment 7. A "✓" indicates that the prediction of a given account is consistent with the results of Experiment 7. A "✗" indicates that it is not consistent with the results. A "()" indicates that it may be consistent, but with some qualifications.

Account	Order	Prediction	Choice	RT
Structural	Head-initial	Asymmetry	✗	✓
locality	Head-final	Asymmetry	✗	✓
Linear	Head-initial	Neutralization[a,b]	(✓)[b]	✗
locality	Head-final	Neutralization[a,b]	(✓)[b]	✗
Frequency	Head-initial	Reversal[c]/Neutralization[d]	(✓)[d]	✗
	Head-final	Reversal[c]/Neutralization[d]	(✓)[d]	✗
Main clause	Head-initial	Neutralization[e]	(✓)[e]	✗
similarity	Head-final	Neutralization[e]	(✓)[e]	✗

[a] is based on counting linear interveners (see footnote 3 of this chapter);

[b] is based on a more liberal application of similarity-based interference;

[c] assumes that the overall prevalence of PV carries over to pronominalized co-arguments;

[d] assumes the English facts about pronominality and ORCs hold true for Tagalog;

[e] assumes the main clause order based on footnote 3.

5.7 Experiment 8: Comparing ambiguous RCs with R-expression co-arguments

This section compared how ambiguous RCs were processed. The verbs exhibited the recent perfective aspect, a grammatical aspect used to denote events that have occurred in the recent past. Recall that this aspect, as discussed in Chapter 4.3.1, exhibits two quirky properties: (i) the verb is standardly analyzed as having no voice morphology; and (ii) both arguments *typically* receive the same case marking (Kroeger, 1993; M. Frazier & Yoshida, 2013). These properties make these sentences globally ambiguous. However, my consultants have indicated a strong preference to parse the first argument as the subject. This is corroborated by the experimental findings of Sauppe (2016) using the Visual World paradigm. An example of a sentence involving the recent perfective is provided in (34).

(34) Kakabati lang ng bata ng mga bisita
greet.RP just GEN child GEN PL visitor
Preferred: 'The child just greeted the visitors.' *V-S-O*
Okay: 'The visitors just greeted the child.' *V-O-S*

However, this ambiguity completely disappears if the one of the co-arguments carries dative, instead of genitive. The argument carrying dative is now obligatorily interpreted as the object, and thus, the argument carrying genitive is now obligatorily the S. This is exemplified in (35).

(35) Kakabati lang ng bata sa mga bisita
greet.RP just GEN child DAT PL visitor
'The child just greeted the visitors.' *V-S-O*
*"The visitors just greeted the child.' **V-O-S*

When relative clauses are involved, the ambiguity found in (34) persists, as in (36-a). The dative case marker also disambiguates in RCs, as in (36-b). I leveraged these properties in this experiment.

(36) a. bata-ng kakabati lang ng mga bisita
child-LNK greet.RP just GEN PL visitor
'Child that just greeted the visitors' *SRC*
'Child that the visitors just greeted' *ORC*

b. bata-ng kakabati lang sa mga bisita
child-LNK greet.RP just GEN PL visitor
'Child that just greeted the visitors' *SRC*
*Child that the visitors just greeted **ORC*

This design allows for the second flavor of the asymmetry to be investigated, as we saw in Dutch (L. Frazier, 1987) and Santiago Laxopa Zapotec (S. Foley et al., 2019): if a language has globally ambiguous RCs, comprehenders offer fewer ORC-interpretations rel-

ative to SRC-interpretations. In a nutshell, I observed the same patterns as in Experiment 6. In unambiguous SRCs, participants, as expected, offered more SRC-interpretations with high levels of accuracy. In ambiguous RCs, they offered more SRC-interpretations and were faster at offering SRC-interpretations.

5.7.1 Materials

The experiment employed a 2×2 design, crossing whether the relative clause was head-initial or head-final (ORDER: HI, HF) and whether the relative clause was ambiguous or not (AMBIG: YES, NO). I created 16 items involving reversible predicates. Each item was distributed evenly across four lists via Latin Square design. The items of this experiment were combined together in one session and randomized with the 72 items from the other experiments. Provided in Table 5.13 is a sample experimental item.

5.7.2 Prediction

Structural locality

Let's continue to assume that the algorithm for counting intervening syntactic projections operates at the level where co-arguments are base-generated. This predicts an asymmetry, where SRCs will be processed easier than ORCs.

Linear locality

Recent perfectives with two genitive arguments are globally ambiguous, suggesting that both VSO and VOS word orders are possible. While my consultants have acknowledged the ambiguity, they have also indicated a strong preference to parse the first argument as the subject. Assuming that this VSO preference is the basis for calculating linear locality, there are predictions to be made. The estimated linear distance in recent perfectives are provided in (37).

Table 5.13: Sample experimental items used in Experiment 8, manipulating ORDER (HI, HF), and AMBIG (YES, NO). The head noun is bolded and the RCs are demarcated by "[]."

ORDER	AMBIG					
		May	isang	palaka	at	kuneho
		EXIST	one.LNK	frog	and	rabbit
		'There is a frog and a rabbit.'				
		Minsan	gusto	nilang	manghuli	
		sometimes	want	3PL.GEN.LNK	maN.catch	
		'Sometimes they like to catch (things).'				
		Minsan	naman	gusto	nilang	magpahuli
		sometimes	also	want	3PL.GEN.LNK	magpa.catch
		'Sometimes they also likes to be caught.'				
		Piliin	ang	larawan	ng …	
		choose.PV	NOM	picture	GEN	
		'Choose the picture of …'				
HI	YES	**palaka**	na	[kakahuli lang	ng	kuneho]
		pig	LNK	caught.RP just	GEN	rabbit
		'…the frog that just caught the rabbit.'				
		'…the frog that the rabbit just caught.'				
HI	NO	**palaka**	na	[kakahuli lang	sa	kuneho]
		frog	LNK	caught.RP just	DAT	rabbit
		'…the frog that just caught the rabbit.'				
HF	YES	[kakahuli lang	ng	kuneho]	na	**palaka**
		caught.RP just	GEN	rabbit	LNK	frog
		'…the frog that just caught the rabbit.'				
		'…the frog that the rabbit just caught.'				
HF	NO	[kakahuli lang	sa	kuneho]	na	**palaka**
		caught.RP just	DAT	rabbit	LNK	frog
		'…the frog that just caught the rabbit.'				

(37) Linear distance in ambiguous recent perfectives

a. Head ① V.RP ___ Object *Linear distance = 1*

b. Head ① V.RP ② Subject ___ *Linear distance = 2*

c. V.RP ___ ① Object Head *Linear distance = 1*

d. V.RP Subject ___ Head *Linear distance = 0*

Accounts emphasizing linear distance would predict an asymmetry in head-initial RCs because the distance between the head noun and the gap is smaller in SRCs than in ORCs. On the other had, they would predict a reversal in head-final RCs since the distance between the head noun and the gap is smaller in ORCs than in SRCs.

Frequency

I do not have frequency data for recent perfectives in matrix clauses or RCs.

Main clause similarity

Recent perfectives with two genitive arguments are globally ambiguous, suggesting that both VSO and VOS word orders are possible. While my consultants have acknowledged the ambiguity, they have also indicated a strong preference to parse the first argument as the subject. Assuming that this VSO preference is the basis for calculating main clause similarity, there are predictions to be made. In head-initial RCs, we would predict an asymmetry. SRCs will be easier to process because the order of SRC reflects the VSO order of main clauses, where S precedes O. In head-final RCs, we would predict a reversal. ORCs will be easier to process because the order of ORC reflects the VSO order of main clauses, where O is after S.

5.7.3 Data analysis

For the choice data, I estimated cumulative link mixed models using the *Ordinal* package in R. I used their derived ordinal ratings as the dependent measure, and the fixed effects were Order (hi, hf), Ambig (yes, no), and their interactions. Order and Ambig were sum-coded, such that hi and no mapped to the negative coefficients. I then included the maximal random effects structure justified by the design that allowed the models to converge.

I also conducted a subset analysis to make sure that the SRC-interpretations in the unambiguous conditions are not attenuating potential differences in the ambiguous conditions. I also conducted a subset analysis of their choice data, looking only at only the ambiguous RCs. I estimated cumulative link mixed models using the *Ordinal* package in R. I used their derived ordinal ratings as the dependent measure, and the fixed effect was Order (hi, hf). I included the maximal random effects structure justified by the design that allowed the models to converge.

For the RTs, I subsetted the data such that I only compared the RTs of the SRC-interpretations of the ambiguous RCs with the RTs of the unambiguous SRCs. I estimated linear mixed effects models using *lme4* in R. I used their log-transformed RT as the dependent measure, and the fixed effects were again Order (hi, hf), Ambig (yes, no), and their interactions. Order and Ambig were again sum-coded, such that hi and no mapped to the negative coefficients. I then included the maximal random effects structure justified by the design that allowed the models to converge.

I also conducted a separate analysis comparing only the RTs of ambiguous RCs to see the difference in how quickly they offer SRC-interpretations compared to ORC-ones. I estimated linear mixed effects models using *lme4* in R. I used their log-transformed RT as the dependent measure, and the fixed effects were again Order (hi, hf), Parse (src, orc), and their interactions. Order and Parse were sum-coded, such that hi and src mapped to the negative coefficients. I then included the maximal random effects structure justified by the design that allowed the models to converge.

5.7.4 Results and discussion

Choice data

In Figure 5.6, I report the breakdown of their responses by interpretation and confidence. In Table 5.14, I provide the summary of the estimated model.

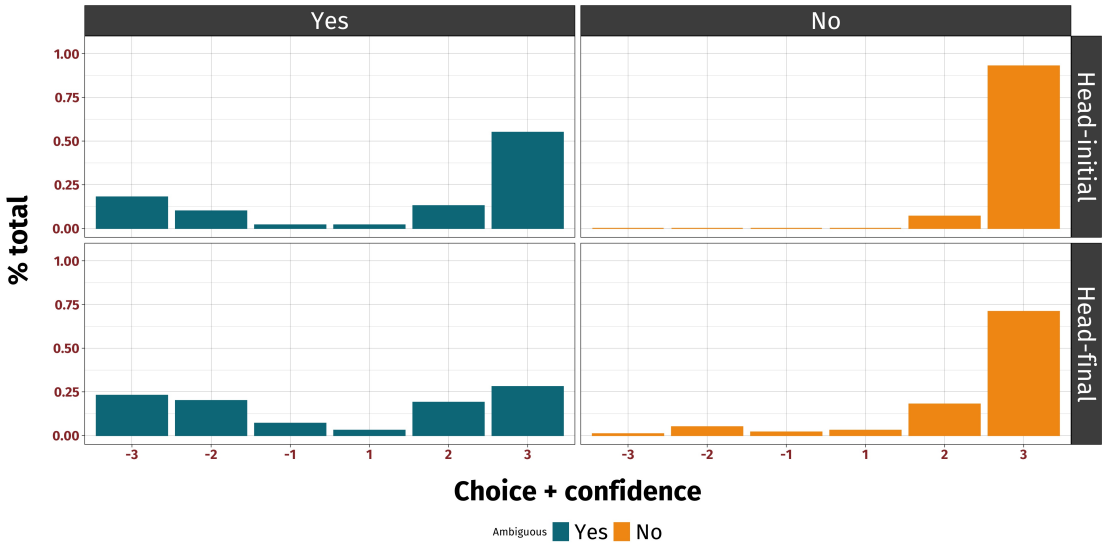


Figure 5.6: Breakdown of participants' responses by interpretation and confidence in Experiment 8. Negative values on the x-axis indicate ORC-responses, while positive values indicate SRC-responses. High values (3) indicate high confidence in their response (irrespective of interpretation), while low values indicate low confidence in their response.

Table 5.14: Summary of cumulative link mixed models in Experiment 8. (Top table) The models included derived choice ratings as the dependent measure, and sum-coded ORDER, AMBIG, and their interactions into the models. (Bottom table) Subset analysis included derived choice ratings as the dependent measure, and sum-coded ORDER as fixed effect.

```
clmm(Rating~Ambig * Order + (1 + Ambig * Order | Subject ) + (1 + Ambig * Order
| Item ))
```

	Estimate	*SE*	z
ORDER	-2.16	.48	-4.48**
AMBIG	-3.38	.50	-6.82**
ORDER×AMBIG	1.60	.92	1.73~

~$p < .1$, *$p < .05$, **$p < .001$

```
clmm(Rating~Order + (1 + Order | Subject ) + (1 + Order | Item ))
```

	Estimate	*SE*	z
ORDER	-1.22	.36	-3.41**

~$p < .1$, *$p < .05$, **$p < .001$

I found a main effect of Order, which suggests that participants were more confidently offering SRC-interpretations in head-initial RCs than in head-final RCs. I also found a main effect of Ambig, which suggests that participants were more confidently offering SRC-interpretations in unambiguous SRCs than ambiguous RCs.

I conducted a subset analysis to make sure that the overwhelming SRC-interpretations in the unambiguous conditions are not attenuating potential differences in the ambiguous conditions. I found that when the unambiguous SRCs were excluded from the analysis, participants were more likely to confidently give SRC-interpretations in head-initial RCs than in head-final RCs. This is consistent with the main effect of Order found in the main analysis.

RT data

I first subsetted the data such that I compared the RTs of the SRC-interpretations of the ambiguous RCs with the RTs of the unambiguous SRCs. I then conducted a separate analysis comparing only ambiguous RCs and how quickly they offer SRC-interpretations relative to ORC-interpretations. In Table 5.15, I provide the summary of the estimated models.

For the first analysis, I found a main effect of Order, which suggests that participants were faster at offering SRC-interpretations in head-initial RCs. I also found a main effect of Ambig, which suggests that participants were faster at offering SRC interpretations in unambiguous SRCs than in ambiguous RCs.

For the second analysis, I found a main effect of Order, which suggests that participants were faster at offering an interpretation when the RC was head-initial. I also found a main effect of Parse, which suggests that participants were faster at offering SRC-interpretations than ORC-interpretations. These effects were qualified by a significant scOrder$\times$Parse interaction. Participants were faster at offering SRC-interpretations

Table 5.15: Summary of linear mixed-effects models in Experiment 8. (Top table) The models included log-transformed RT as the dependent measure, and sum-coded ORDER, AMBIG, and their interactions into the models. (Bottom table) Ambiguous only. The models included log-transformed RT as the dependent measure, and sum-coded ORDER, PARSE, and their interactions into the models.

```
lmer(log(RT)~Ambig * Order + (1 + Order | Subject ) + (1 | Item ))
```

	Estimate	*SE*	z
Intercept	7.01	0.05	140.740**
ORDER	.52	.05	9.96**
AMBIG	.17	.04	4.64**
ORDER×AMBIG	.13	.07	1.75 ~

~$p < .1$, *$p < .05$, **$p < .001$

```
lmer(log(RT)~Parse * Order + (1 | Subject ) + (1 | Item ))
```

	Estimate	*SE*	z
Intercept	7.17	.05	156.69**
ORDER	.38	.04	8.68**
PARSE	.19	.05	3.93**
ORDER×PARSE	-.37	.09	-3.9**

~$p < .1$, *$p < .05$, **$p < .001$

than ORC-interpretations in head-initial RCs compared to in head-final RCs. Pairwise comparison of SRC-interpretations and ORC-interpretations in head-initial RCs indicate that they were faster offering an SRC-interpretation than an ORC-one ($p < .001$). Pairwise comparison of SRC-interpretations and ORC-interpretations in head-final RCs did not reveal any differences ($p = .89$).

Discussion

Analyses of the choice data suggest an asymmetry when the RC was head-initial. There is no evidence for an asymmetry when the RC was head-final. Anaalyses of the RT data suggest once again suggest an asymmetry when the RC was head-initial. There

is no evidence for an asymmetry when the RC was head-final.

How does each account fare? In Table 5.16, I provide an overview of how each account fared. A "✓" indicates that the prediction of a given account is consistent with the results, while a "✗" indicates that it is not consistent. A "()" indicates that it may be consistent, but with some qualifications.

Table 5.16: An evaluation of how each account fares compared against the results of Experiment 8. A "✓" indicates that the prediction of a given account is consistent with the results of Experiment 8. A "✗" indicates that it is not consistent with the results. A "()" indicates that it may be consistent, but with some qualifications.

Account	Order	Prediction	Choice	RT
Structural locality	Head-initial	Asymmetry	✓	✓
	Head-final	Asymmetry	✗	✗
Linear locality	Head-initial	Asymmetry	✓	✓
	Head-final	Reversal	✗	✗
Frequency	Head-initial	Unclear	–	–
	Head-final	Unclear	–	–
Main clause similarity	Head-initial	Asymmetry	✓	✓
	Head-final	Reversal	✗	✗

Overall, accounts emphasizing structural locality fared relatively well. Accounts emphasizing linear locality and/or similarity to main clauses also fared relatively well. None of these proposals predicted the neutralization in the participants' choice data and RT data in head-final RCs.

5.8 General discussion

In this chapter, I asked the following questions. First, are SRCs easier to process than ORCs in Tagalog? Second, if they are, can other factors can modulate this asymmetry? Lastly, what can the empirical landscape of relative clause processing in Tagalog tell us about the classes of explanations that have been proposed in the literature to account for

what seems to be a robust asymmetry cross-linguistically?

In Experiment 6, participants were more accurate and confident in SRCs than in ORCs when the RC was head-initial. Meanwhile, they were as accurate and confident in SRCs as they were in ORCs when the RC was head-final. Overall, they were faster at providing correct SRC-interpretations than ORC-interpretations. These results add to the body of literature showing that Tagalog exhibits the classic subject-object asymmetry, replicating the earlier findings about Tagalog head-initial RCs (Pizarro-Guevara, 2014; Bondoc et al., 2018; Tanaka et al., 2019). These results also add to the growing body of literature showing that word order does attenuate the asymmetry, echoing what Wagers et al. (2018) found in Chamorro.

In Experiment 7, when the co-arguments inside the RC were pronominalized, there is no evidence for asymmetry in their choice data in both word orders. These results, coupled with the results of Experiment 6, add to the body of literature showing that the referential type of intervening elements attenuate the asymmetry, echoing what P. C. Gordon et al. (2004) found in English and what Price and Witzel (2017) found in Russian. The persistence of the asymmetry in their RT data provides strong evidence for the universality of this subject/object asymmetry.

In Experiment 8, when the RCs were ambiguous, participants offered more SRC-interpretations than ORC-interpretations when the RC was head-initial. Again, this pattern of results suggest an asymmetry, and are akin to what L. Frazier (1987) found in Dutch and what S. Foley et al. (2019) found in Santiago Laxopa Zapotec. The attenuating effects of word order echo the results of Experiment 6 with unambiguous RCs.

To recap, while the findings of each experiment were different, a recurring theme is that in places where one saw attenuation in their choice data—either via word order or pronominality of the co-argument, the asymmetry persisted in their RT-data.

What do these patterns of results tell us about the classes of explanations? Framed

differently, which proposal(s) fared best? My goal here was never to argue for a single account to explain all of the results. Language comprehension involves coordinating multiple sources of information. There was not a single account that captured all of the contours of the data. In places where one account could explain the asymmetry in one condition/measure, said account is unable to explain the attenuation in another condition/measure. For example in Experiment 6, when the RCs involved were head-final, structure-based accounts fared well in capturing the asymmetry in the participants' RT data, but not very well in capturing the attenuation in their choice data. On the other hand, accounts emphasizing main clause similarity fared well in capturing the attenuation in their choice data, but not very well in capturing the asymmetry in their RT data. In Experiment 7, structure-based accounts again fared well in capturing the asymmetry in their RT data; the other accounts fared well in capturing the neutralization in their choice data.

Lastly, I acknowledge that I might not have given frequency-based accounts their best chance to succeed. Recall that Nagaya (2019) found that ORCs are more common than SRCs in the language, irrespective of word order. Pure frequency-based accounts would predict a reversal of the asymmetry. None of the measures indicated a reversal. In Experiment 6, I offered some potential explanations as to why their predictions might be misrepresented. Here, I offer another potential explanation: timing of when the difficulty is expected to arise.

Staub (2010) argued that frequency-based accounts predicted earlier difficulty, at the relative clause subject, in English. The reasoning is as follows: because SRCs are more frequent in English, after encountering the complementizer *that*, an SRC is a more likely continuation. When we see an overt subject, we are surprised and this causes processing difficulty. In an eye-tracking while reading study, Staub found that regressive saccades were much more likely from the subject DP of an ORC. He interpreted this as being con-

sistent with the predictions of frequency-based accounts. He also found longer reading times on the verb in ORCs. He interpreted this finding as being consistent with the predictions of memory-based accounts, another class of account that he was considering. Taken together, he maintained that both accounts contributed to the processing difficulty of ORCs. Crucially, they differed qualitatively, in terms of when the difficulty arose.

We could use the same reasoning for Tagalog RCs. Because ORCs are more frequent in Tagalog, after encountering the linker in a head-initial RC, an ORC continuation is a more likely continuation. When we see an AV morphology on the verb, we are surprised and this could cause processing difficulty. In order to evaluate whether there is an *early* reversal in Tagalog RC processing, we need to use a measure that allows moment-by-moment analyses—like eye movement in a simplified visual world paradigm, for instance—instead of analyzing only measures that indicate final states of comprehension (e.g., picture choice and RT). In other words, perhaps we are not seeing the predicted reversals because we are not looking for them early enough.

Chapter 6

Conclusion

This dissertation is part of a broader trend in psycholinguistics that incorporates data from "smaller languages" to address the WEIRD-centricity that the field suffers from. The empirical focus of this dissertation has been how Tagalog comprehenders understand FGDs in the language. I reported the results of a series of acceptability judgment studies, and stops-making-sense and picture-matching experiments on the comprehension of various FGDs in Tagalog.

This dissertation has three main take-aways. First, the extraction restriction in Tagalog is not clean-cut for some speakers. This finding might be most relevant to Austronesianists more generally and generative syntacticians, more specifically. Second, voice is a cue that Tagalog comprehenders use when processing FGDs. The way in which voice was leveraged in real-time seemed to be mediated by the type of dependency involved. Third, the subject-object asymmetry is quite robust in Tagalog. SRCs are easier to process than ORCs in Tagalog. The relative order of the head noun and the RC, as well as the referential type of co-arguments, can modulate this asymmetry, however. The last two findings might be most relevant to psycholinguists and acquisitionists.

6.1 The extraction restriction is not clean-cut

Voice has been central to many investigations in Austronesian syntax because it interacts with case marking, verbal morphology, and A-bar movement. Here I looked at the interaction between voice and A-bar movement in Tagalog. In Chapter 2, we saw that the received generalization is that the extraction restriction is clean-cut: only the extracted argument that is cross-referenced by voice morphology is eligible for to participate in filler-gap dependencies (Schachter, 1977; Ceña, 1979; Aldridge, 2002; Rackowski & Richards, 2005; Kaufman, 2009; Law, 2016; Aldridge, 2017, 2018). When the verb exhibits AV, only the subject (=agent) is extractable. When the verb exhibits PV, only the object (=patient) is extractable. When the verb exhibits the various flavors of applicative morphology, only the applied object is extractable.

In Chapter 3, I showed that some speakers productively allowed the extraction of subjects/agents irrespective of voice. I am not claiming that I am the first one to notice this variation. The earliest documentation of this variation that I could find in the generative syntax literature is in the work of de Guzman (1995)—modulo sub-extraction, which was observed even earlier by Ceña (1979). Ceña and Nolasco (2011, 2012) had also observed this variation but these works were in Tagalog. Unfortunately, this has remained undiscussed in the mainstream generative syntax literature—even in works that develop accounts for the extraction restriction in the language. In Chapter 3 of my dissertation, I did four things. First, I made available the empirical generalizations of Ceña and Nolasco (2011, 2012) available in English. I also provided (i) naturally occurring data and (ii) experimental and computational evidence that are consistent with an extraction restriction that is not as clean-cut. Lastly, I conjectured that this variation can be recast as grammar competition in learning in the vein of (Han et al., 2007). When the extracted argument is cross-referenced by voice, the linguistic output of these grammars converge. Their output diverge only when we look at the speakers' behavior when the extracted argument is

not cross-referenced by voice.

6.2 Voice is a cue used in FGD-processing

As we saw in Chapter 4, there is a wealth of evidence that comprehenders do not wait until they encounter all of the bottom-up information in the input. Instead, they use various types of linguistic information to predictively posit a gap that would allow the dependency to be resolved. They can use various classes of linguistic information, like syntactic (Traxler & Pickering, 1996), lexical (Trueswell et al., 1994), morphological (Kamide et al., 2003), and prosodic (Nagel et al., 1994) information.

I examined whether Tagalog comprehenders use the language's voice morphology to guide their incremental interpretations. I hypothesized that voice allows comprehenders to commit to an interpretation upon encountering the verb, since they have at this point in time information about the event structure and by virtue of the voice morphology, the thematic role of the filler. In Experiments 4 and 5, I used the stops-making-sense paradigm (Boland et al., 1990, 1995) to ask whether and how voice is asked as a cue in FGD-processing. I found that Tagalog comprehenders used voice to reliably link the filler to the gap and commit to the correct interpretation before encountering disambiguating information in the form of the co-argument. Thus, I added Tagalog to the general picture that comprehenders actively link the filler to the gap and they can use morphological information, like voice morphology, to do so. I also found that the way in which comprehenders leveraged voice as a cue was mediated by the type of filler-gap dependencies.

I related the results of Experiments 1, presented in Chapter 3, to the results of Experiment 5 to talk about the level of granularity of the information about voice that comprehenders use in real-time. I argued that comprehenders were using finer-grained information about voice when processing FGDs. However, the results suggest that they also engaged with other classes of linguistic information, including (but not limited to) infor-

mation about the structural similarities and the thematic complexity of the dependencies involved, and the relative frequency of the different types of voices in the language. These interactions resulted in processing asymmetries.

6.3 SRCs are easier to process than ORCs in Tagalog

One of the common findings in RC processing research is that comprehenders prefer SRCs over ORCs. This asymmetry comes in two flavors. First, in languages with unambiguous relative clauses, ORCs have been found to be more difficult to process than SRCs (Staub, 2010). Second, in languages with ambiguous relative clauses, comprehenders offer more SRC-interpretations than ORC-interpretations (L. Frazier, 1987).

I leveraged the fact that Tagalog has both unambiguous and ambiguous RCs and investigated how SRCs and ORCs were processed in the language. Across three experiments, I found that SRCs were easier to process than ORCs, which echoes previous findings on Tagalog RC-processing (Pizarro-Guevara, 2014; Bondoc et al., 2018; Tanaka et al., 2019). I also leveraged the fact that the language can have both head-initial and head-final RCs—a typologically rare property for a language to have (Dryer, 2013)—to see how this could affect the asymmetry. The relative order of the head noun and the RC did have attenuate the asymmetry, such that head-final RCs have consistently shown a reduction—and sometimes even the neutralization—of the asymmetry. These findings echo the findings in Chamorro (Wagers et al., 2018). The findings of Experiments 6 and 7 provide evidence that the referential type of the co-argument can also attenuate the asymmetry, consistent with the findings in English (P. C. Gordon et al., 2004).

There has not been a shortage of proposals advanced in the literature to explain this asymmetry. However, no one view can fully capture the texture of the Tagalog data. Is this alarming? Not one bit. Perhaps a reasonable way to proceed would be to stop treating these accounts separately and argue for an integrative account where the asymme-

try is a composite phenomenon. After all, language comprehension is a complex phenomenon that involves coordinating multiple sources of information—both linguistic and non-linguistic. This sentiment is perhaps best captured by Wagers et al. (2018, p. 208):

> Each of these accounts—a universalist account grounded in typological observations, a memory-based account that trades on the difficulty of forming dependencies between non-adjacent constituents, and an experience-based account linked to the distributions of constructions, rules, or substrings—is deeply correlated with the others. And they may all simultaneously be correct descriptions of some linguistic process at some level of analysis.

If we accept that these different classes of information all contribute in some way to the asymmetry, then the question becomes one of "when": When do they contribute to the difficulty? How are they temporally prioritized with respect to one another?

6.4 Future directions and concluding remarks

This dissertation leaves open many questions. Below I make explicit some of these questions and will talk about how I could address them in future studies. The overarching goal of these studies is to have a comprehensive understanding of the empirical terrain with regards to movement dependencies in Tagalog to further inform and enrich theories in both formal linguistics and psycholinguistics.

6.4.1 Two grammars of Tagalog?: Collecting child data

In Chapter 3, I had speculated that the competing generalizations of the extraction restriction in Tagalog could be a product of grammar competition in learning à la Han et al. (2007). Recall that the main intuition behind grammar competition in learning is that sometimes the primary linguistic data that a child in a language L is exposed to are compatible with at least two hypotheses, H_1 and H_2. When linguistic experience cannot adjudicate between the two, some learners acquire a grammar based on H_1 and others, one

based on H_2. If we take seriously this reformulation, then we should expect to find some evidence of these two grammars in the how children judge voice mismatch extractions.

To be able to model the children's judgment process of voice mismatch extractions, similar to the ones presented in Chapter 3, we would need their judgment ratings. Working with children is not an easy feat; it requires imagination because these "participants" get bored easily. To get their ratings, I will use a slightly modified reward/punishment version of a grammaticality judgment task (P. Gordon, 1983; Hochberg, 1986)[1]. This methodology typically involves introducing children to puppets in a doll house. Children are told that puppets will sometimes will say something that sounds "okay" and sometimes, "silly." They give a gold star to the puppet who had said it okay.

For my adapted version, I am thinking of pre-recording sentences that involve voice (mis)match extractions, akin to the items in Experiments 1–3, and presenting these auditorily. This mode of presentation preemptively removes the child participants' level of literacy as a potential confound. Children will be given a tablet so they can play a "game." Framing them as such might make these experiments more palatable. They will be told that there will be cute animals, the ones used in Experiments 6–8, and these animals will be saying sentences that sometimes sound like a good sentence in Tagalog, and sometimes, silly sentences that don't even sound like Tagalog. They will be told that these animals really enjoy eating sweet candies and really dislike eating sour candies. If an animal says a sentence that sounds like a good sentence in Tagalog, the children are supposed to give the animal a sweet candy as a reward. If the animal says a sentence that sounds silly in Tagalog, they are supposed to give the animal a sour candy as a punishment/a signal to the animal that it can do better.

To get more texture out of this binary response, for each trial, the children will also be asked to give their "confidence rating." Confidence rating will take the form of asking

[1] Reward/punishment are used extensively in truth-value judgment tasks (Crain, 1991, *et seq.*), a slightly different type of binary-response task. In a typical TVJ trial, children hear a sentence from a sock puppet

them how many sweet or sour candies they want to give this animal. They will be instructed beforehand to give 3 sweet candies if the sentence sounds "really really good," 2 sweet candies if the sentence sounds "good," and 1 sweet candy if the sentence is "good, but not that good." Similarly, they will be instructed beforehand to give 3 sour candies if the sentence sounds "really really silly," 2 sour candies if the sentence sounds "silly," and 1 sour candy if the sentence is "silly, but not that silly." These "confidence ratings" will allow me to derive an ordinal scale of their ratings, which can then be used to estimate the underlying distribution of their judgment process. I provide in Figure 6.1 how a typical trial will look using this adapted reward/punishment judgment study.

6.4.2 Do different FGDs exhibit varying mismatch penalties?

Recall that in Experiments 2 and 3, I found that voice mismatch extractions incurred a smaller mismatch penalty in non-AV voices when the subject/agent is extracted. This is consistent with the results of Experiment 1. However, unlike Experiment 1, there was no evidence that relative clauses exhibited greater asymmetry than *wh*-questions in Experiments 2 and 3. It could be that what I found in Experiment 1 was a fluke and that there is no real difference in how permissive *wh*-questions and relative clauses are with respect to voice mismatch extractions. However, I am committed to the differences shown in Experiment 1 based on my own work with speakers. Thus, an open question is why these differences were not observed here in Experiments 2 and 3.

I conjectured that the underlying cause for difference is the slight differences in the experimental items used. These non-converging findings might be artifactual. Recall that in Experiment 1, I used matrix *d*-linked *wh*-questions, while in Experiments 2 and 3, I used embedded bare *wh*-questions. One obvious difference is that the first one involved *d*-linked questions, while the last two involved bare interrogatives. Another obvious difference is that the first one involved matrix questions, while the last two involved em-

Figure 6.1: Sample item of the proposed reward/punishment judgment study for children. The top panel represents the screen where the child participants hear the experimental item, which could either be voice match or voice mismatch extraction. The middle panel represents the screen where the child participants reward or punish the animal by giving him a sweet or a sour candy. The bottom panel represents the "confidence rating" where they get to choose how much of a reward or a punishment they will give the animal.

bedded questions. Thus, the inconsistency could be about the difference in the type of interrogative used, or the embedding context, or both!

In order to isolate the underlying cause for the inconsistencies, I intend to run four

experiments simultaneously, two involving matrix questions and two involving embedded questions. For the 2 experiments that will be comparing mismatch penalties under AV and PV (akin to Experiments 1 and 2), both will use a $2\times2\times2$ factorial design, crossing whether the verb exhibited AV or PV (VOICE: AV, PV), whether the extracted argument is cross-referenced by voice or not (MATCH: MATCH, MISMATCH), and whether the interrogative used is bare or *d*-linked (TYPE: BARE, DLINKED). For the 2 experiments that will be comparing mismatch penalties under applicatives (akin to Experiment 3), both will use a $2\times2\times2$ factorial design, whether the extracted argument is cross-referenced by voice or not (MATCH: MATCH, MISMATCH), whether the extracted argument is the subject/agent or not (AGEX: YES, NO), whether the interrogative used is bare or *d*-linked (TYPE: BARE, DLINKED). In other words, their designs will closely mirror the design of the experiments reported in Chapter 3, but with one additional manipulation—the type of interrogative used. To remove potential animacy effects in *d*-linking, the appropriate and universally accepted *d*-linked interrogative frames will be used: *aling N* for inanimates and *sinong N* for animates. By running half is embedded questions and half as matrix questions, we would be able to see if embedding contexts factors in the inconsistencies. Provided in Table 6.1 are sample experimental items involving only embedded questions under AV and PV. For their matrix counterparts, simply remove "Pedro asked."

6.4.3 Voice mismatches in islands

In Chapter 3, I had claimed that (i) Tagalog is an island-sensitive language, independent of voice mismatches; and (ii) for speaker who allow voice mismatches, extraction of the subject/agent under non-AV voices also exhibits the same sensitivity to islands. I used island-sensitivity as the main piece of evidence for arguing a movement-based derivation of these voice mismatch extractions.

Table 6.1: Sample experimental items used to isolate the underlying cause for the inconsistencies. These are modified versions of the experimental items used in Experiment 2, manipulating VOICE (AV, PV), MATCH (+, –), and TYPE: BARE, DLINKED.

TYPE	VOICE	MATCH						
			Tinanong	ni Pedro	kung …			
			asked.PV	GEN PN	COMP…			
			Pedro asked…					
WHQ	AV	+	sino/sinong sakristan	ang	kumakanta	ng himno	para sa madre	gabi-gabi.
			who/which altar boy	NOM	sing.AV	GEN hymn	for DAT nun	every night.
			…who/which altar boy sings a hymn for the nun every night.					
	AV	–	ano/aling himno	ang	kumakanta	ang sakristan	para sa madre	gabi-gabi.
			what/which hymn	NOM	sing.AV	NOM altar boy	for DAT nun	every night.
			…what the altar boy sings for the nun every night.					
	PV	+	ano/aling himno	ang	kinakanta	ng sakristan	para sa madre	gabi-gabi.
			what/which hymn	NOM	sing.PV	GEN altar boy	for DAT nun	every night.
			…what the altar boy sings for the nun every night.					
	PV	–	sino/sinong sakristan	ang	kinakanta	ang himno	para sa madre	gabi-gabi.
			who/which altar boy	NOM	sing.PV	NOM hymn	for DAT nun	every night.
			…who/which altar boy sings the hymn for the nun every night.					

I offered two types of evidence to show that Tagalog is island-sensitive—fieldwork data and the results of acceptability judgment studies investigating *wh*-islands and complex NP islands (i.e., the filler items of Experiments 2 and 3). The experimental items used in said judgment studies only involved voiced match extractions, and used a 2×2 design, crossing whether the verb exhibited AV or PV (VOICE: AV, PV) and whether the extraction site is an island or not (ISLAND: NO, YES). I had not done any systematic investigation of voice mismatch extractions in island contexts, but a follow-up study comparing voice match and voice mismatch extractions in island-contexts would simply require an extra experimental manipulation. These judgment studies would use a 2×2×2 design, crossing whether the verb exhibited AV or PV (VOICE: AV, PV), whether the extraction site is an island or not (ISLAND: NO, YES), and crucially, whether the extracted argument is cross-referenced by voice or not (MATCH: MATCH, MISMATCH).

6.4.4 What other cues do Tagalog comprehenders use?

In Chapter 4, I investigated the role of voice morphology in processing filler-gap dependencies. Across two experiments, I showed that this piece of morphological information allows Tagalog comprehenders to actively associate the filler and the gap in real-time. Recall that it is always an empirical question what cues Tagalog comprehenders could be attending to. There is evidence that prosody in English is a rich source of information that allows comprehenders to anticipate an upcoming referent, prior to encountering disambiguating lexical information (Snedeker & Trueswell, 2003; Ito & Speer, 2008).

Using a cross-modal lexical priming task, Nagel et al. (1994) argues that English that comprehenders are able to infer the location of the gap based on the prosodic contour of sentences like those provided in (1). At the direct object position, they used the following lexical probes: (i) a semantic associate of the filler (e.g., *patient*); (ii) an unrelated control (e.g., *current*); and (iii) a orthotactic-obeying nonce word (e.g., *comtent*). They found prim-

ing effects for the semantic associate at the direct object position in (1-a) but not in what could have been the direct object in (1-b).

(1) Adapted from (Nagel et al., 1994)

a. Which doctor did the supervisor call to get help for his youngest daughter?

b. Which doctor did the supervisor call to get help for during the crisis?

In order to investigate the role of prosodic cues in FGD-processing, there is crucial legwork that needs to be done in order to gain a better understanding of the prosodic system of the language. Unfortunately, there is very little work on Tagalog prosody. I know of three studies: Hsieh (2016), Richards (2017), and Nagaya and Hwang (2018). Note that these studies only describe prosody in what I refer to as run-of-the-mill declaratives. To my knowledge, there has not been any studies that describe prosody in Tagalog FGDs.

Let's consider Richards (2017) as a case study to speculate how a comprehender might use prosodic cues to process FGDs. His main finding is that "the object generally has a higher pitch than the subject would in the same position; immediately postverbal objects are higher-pitched than immediately postverbal subjects, and the same is true for objects that are not immediately postverbal." (Richards, 2017, p. 20). Let us assume that his generalization extends to FGDs. Whether Tagalog comprehenders actually attend to prosodic cues is an empirical question. Since this section is already speculative in nature, let's further assume that they do. With all these assumptions in place, a reasonable hypothesis could be that upon encountering the filler, a Tagalog comprehender could determine whether it is a subject or an object based on the relative pitch with which it was produced.

In a Visual World experiment, consider a display that has a baby, a hamster, and an adult.[2] When comprehenders hear the filler *sanggol* 'infant' produced with a relatively

[2]I am further assuming that participants were exposed to the range of pitches of the talker during practice and also during the burn-in trials, which are trials that are placed in the beginning of the experiment that have the same shape as the experimental trials. These trials will not be analyzed, however. The main reason

higher pitch, then they might expect that this filler is associated with an object gap. Upon encountering the verb "carry," there might be more anticipatory looks to what could potentially carry an infant (i.e., the "adult"). On the other hand, if the filler was not produced with a relatively higher pitch, then comprehenders might expect that this filler is associated with a subject gap. Upon encountering the verb "carry," there might be more anticipatory looks to what an infant could potentially carry (i.e., the "hamster").

Again I want to emphasize that this section is speculative. A better understanding of how the prosodic system of Tagalog behaves in both FGD- and non-FGD-contexts is an essential first step for any investigation on the role of prosodic cues in FGD-processing.

6.4.5 Quantifying how informative voice is

In Chapter 4, I related their behavior in an SMS task (Experiments 4–5) to the results of Experiment 1 in Chapter 3. I fully acknowledge that using the judgment study in Experiment 1 is at best, an imperfect proxy for estimating how informative voice is as a cue that signals the identity of the gap given a particular form of voice. It would have been ideal to quantify these values by performing a corpus analysis, in the same vein as the first study of Garcia et al. (2019). Unlike their corpus analysis, this analysis would have involved adult language and more importantly, different FGDs. In order to use a corpus to calculate these values, the corpus would need to be part-of-speech tagged, as well as be dependency parsed. Unfortunately, I do not know of any corpora that satisfy these requirements.

Creating a part-of-speech tagged and dependency-parsed corpus of Tagalog is a big undertaking and this has been a part of my long-term research goals. I do not have all of the necessary skills to actualize this project on my own, so I see this project as an avenue to collaborate with the following people/institutions: (i) Henrison Hsieh, who already has

for including burn-in trials is to give participants a preview of the entire range of possible experimental manipulations, giving them an idea of what counts as relatively higher pitch for this particular talker.

written a very basic morphological parser that can help semi-automate Tagalog sentential data; (ii) Shirley Dita and her team at De La Salle University, who already have the experience compiling a corpus called the Palito corpus (Dita & Roxas, 2011), which functioned mostly as a concordancer and is now defunct unfortunately; and (iii) the Department of Linguistics at the University of the Philippines - Diliman, who is already collaborating with various departments in the UP system that have the necessary skills to actualize this project.

6.4.6 Eye-movement in relative clause processing

In Chapter 5, I had discussed how pure frequency-based accounts predict a reversal but none of the measures that I looked at indicated a reversal. One potential explanation for this is that we might not be seeing the predicted reversal because we are not looking for this effect early enough.

According to Staub (2010), frequency-based accounts predict early difficulty in English, namely as early as the subject of the RC. His reasoning was as follows: because SRCs are more frequent in English, after encountering the complementizer *that*, an SRC is a more likely continuation. When we see an overt subject, we are surprised and this causes processing difficulty.

We could use the same type of reasoning in Tagalog. Recall that Nagaya (2019) found that ORCs are more frequent than SRCs. After encountering the linker in a head-initial RC, an ORC continuation is more likely. When they see AV morphology on the verb, Tagalog comprehenders are surprised and this could cause processing difficulty. When they see PV morphology on the verb, this accords with their expectations and thus, this might not incur as much processing difficulty.

In order to evaluate whether there is an *early* reversal in Tagalog RC processing, we need to use a measure that allows moment-by-moment analyses—like eye movement in a

Figure 6.2: Sample picture selection screen used in the experiments presented in Chapter 5

simplified visual world paradigm, for instance—instead of analyzing only measures that indicate final states of comprehension (e.g., picture choice and RT). In a Visual World paradigm, a reversal means that upon encountering the head noun, the picture where the head noun is the object of the verb will get more looks than the picture where the head noun is the subject of the verb. Consider the picture selection screen provided in Figure 6.2.

In Tagalog, upon encountering the filler *kuneho* "rabbit" and the linker *na*, a reversal would mean more looks to the picture on the right, where the rabbit is being spanked by the monkey, compared to the picture on the left, where the rabbit is spanking the monkey. I do have eye movement data that are not reported in this dissertation. I am still in the middle of processing them, however.

Now let's suppose that after I have analyzed the participants' eye movement data, I did not find the predicted reversal at the verb. Should I throw in the towel and abandon giving frequency-based explanations their best chance to succeed? No, not necessarily.

One could argue that the inconsistency could be due to the fact that Staub (2010) and I used different methodologies. My experiments in Chapter 5 used a paradigm that is more akin to Visual World, while Staub used eye-tracking while reading. They are two inherently different methodologies. The hypothesis I sketched above assumes that the results of eye-tracking while reading port over nicely to the results of Visual World. That may not necessarily be the case.

A more direct comparison can be made if I adapt my experimental items and run a series of eye-tracking while reading studies. This extension is in theory feasible since speakers of Tagalog are literate *and* are used to reading in Tagalog. Furthermore, participant recruitment is not an issue logistically since I already have a good relationship with the University of the Philippines – Diliman. The access to the necessary hardware in the fieldsite is the main factor that limits the feasibility of running these studies.

References

Aldridge, E. (2002). Nominalization and *wh*-movement in Seediq and Tagalog. *Language and Linguistics*, *3*(2), 393–426.

Aldridge, E. (2004a). Antipassive and specificity in tagalog. In P. Law (Ed.), *Proceedings of the 11th Annual Meeting of Austronesian Formal Linguistics Association* (pp. 1–14). Berlin, Germany: Zentrum fur Allgemeine Sprachwissenschaft, Typologie, und Universalienforschung.

Aldridge, E. (2004b). *Ergativity and word order in Austronesian languages* (Unpublished doctoral dissertation). Cornell University, Ithaca, NY.

Aldridge, E. (2008). Phase-based account of extraction in Indonesian. *Lingua*, *118*(10), 1440–69.

Aldridge, E. (2012). Antipassive and ergativity in Tagalog. *Lingua*, *122*(3), 192–203.

Aldridge, E. (2013). Wh-clefts and verb-initial word order in Austronesian languages. In *Cleft structures* (pp. 71–96). Amsterdam: John Benjamins.

Aldridge, E. (2017). Internally and externally headed relative clauses in Tagalog. *Glossa: A journal of general linguistics*, *2*(1), 1–33.

Aldridge, E. (2018). ϕ-feature competition: A unified approach to the Austronesian extraction restriction. In *The Proceedings of the 52nd Meeting of the Chicago Linguistics Society.* Chicago, IL: Chicago Linguistics Society.

Anand, P., Chung, S., & Wagers, M. (2011). Widening the net: Challenges for gathering linguistic data in the digital age. In *Sbe 2020: Future research in the social, behavioral & economic sciences* (p. 121). National Science Foundation.

Aoshima, S., Phillips, C., & Weinberg, A. (2004). Processing filler-gap dependencies in a head-final language. *Journal of memory and language*, *51*(1), 23–54.

Arnon, I. (2010). Rethinking child difficulty: The effect of NP type on children's processing of relative clauses in Hebrew. *Journal of Child Language*, *37*(1), 27–57.

Arosio, F., Guasti, M. T., & Stucchi, N. (2011). Disambiguating information and memory resources in children's processing of Italian relative clauses. *Journal of Psycholinguistic Research*, *40*, 137–54.

Bader, M., & Meng, M. (1999). Subject-object ambiguities in German embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research*, *28*, 121–43.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*, 328. Retrieved from 10.3389/fpsyg.2013.00328

Barr, D. J., Gann, T. M., & Pierce, R. S. (2011). Anticipatory baseline effects and information integration in visual world studies. *Acta psychologica*, *137*(2), 201–7.

Barrios, J. (2014). *Tagalog for beginners: An introduction to Filipino, the national language of the Philippines*. North Clarendon, VT: Tuttle.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Betancort, M., Carreiras, M., & Sturt, P. (2009). The processing of subject and object relative clauses in Spanish: An eye-tracking study. *The Quarterly Journal of Experimental Psychology*, *62*(10), 1915–29.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and development of language* (pp. 279–362). New York, NY: Wiley.

Bickel, B., Witzlack-Makarevich, A., Choudhary, K. K., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2015). The neurophysiology of language processing shapes the evolution of grammar: Evidence from case marking. *PLoS One*, *10*(8), e0132819. doi: 10.1371/journal.pone.0132819

Bloomfield, L. (1917). *Tagalog texts with grammatical analysis*. Urbana-Champaign, IL: University of Illinois.

Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language*, *29*(4), 413–32.

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, *34*(6), 774–806.

Bondoc, I. P., O'Grady, W., Deen, K., & Tanaka, N. (2018). Agrammatism in tagalog: voice and relativisation. *Aphasiology*, *32*(5), 598–617.

Bondoc, I. P., & Schafer, A. (2019). *The role of Tagalog voice information on sentential argument order in production and gaze patterns in comprehension.* Poster presented at the 32nd Annual CUNY Conference on Human Sentence Processing. Boulder, CO.

Brysbaert, M., & Mitchell, D. C. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. *The Quarterly Journal of Experimental Psychology*, *49A*(3), 664–95.

Bulut, T., Cheng, S.-K., Xu, K.-Y., Hung, D. L., & Wu, D. H. (2018). Is there a processing preference for object relative clauses in Chinese? evidence from ERPs. *Frontiers in Psychology*, *9*, 995. doi: 10.3389/fpsyg.2018.00995

Carreiras, M., Duñabeitia, J. A., Vergara, M., de la Cruz-Pavía, I., & Laka, I. (2010). Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition*, *115*(1), 79–92.

Ceña, R. M. (1979). Tagalog counterexamples to the Accessibility Hierarchy. *Studies in Philippine Linguistics*, *3*(1), 119–24.

Ceña, R. M., & Nolasco, R. M. D. (2011). *Gramatikang Filipino: Balangkasan*. Quezon City, Manila: The University of the Philippines Press.

Ceña, R. M., & Nolasco, R. M. D. (2012). *Sintaks ng Filipino*. Manila, Philippines: The National Commission for Culture and the Arts.

Chacón, D. A., Imtiaz, M., Dasgupta, S., Murshed, S. M., Dan, M., & Phillips, C. (2016). Locality and word order in active dependency formation in bangla. *Frontiers in Psychology*, *7*. Retrieved from 10.3389/fpsyg.2016.01235

Chen, B., Ning, A., Bi, H., & Dunlap, S. (2008). Chinese subject-relative clauses are more difficult to process than the object-relative clauses. *Acta Psychologica*, *129*(1), 61–5.

Chen, V. (2017). *A re-examination of the Philippine-type voice system and its implications for Austronesian primary-level subgrouping* (Unpublished doctoral dissertation). University of Hawai'i at Manoa, Honolulu, HI.

Chen, V., & Fukuda, S. (2016). Raising to object out of cp as embedded left dislocations: evidence from three formosan languages. In K. min Kim et al. (Eds.), *Proceedings of the 33rd West Coast Conference on Formal Linguistics.* Sommerville, MA: Cascadilla Proceedings Project.

Chen, V., & McDonnell, B. (2019). Western Austronesian voice. *Annual Review of Linguistics*, *5*, 173–95.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik* (pp. 89–155). Cambridge: MIT Press.

Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 1–52). Cambridge: MIT Press.

Christensen, R. H. B. (2018). *Ordinal—regression models for ordinal data. (version 2018.8-25) [*Computer program*].* Retrieved from `http://www.cran.r-project.org/package=ordinal/`

Christianson, K., & Cho, H. Y. (2009). Interpreting null pronouns (*pro*) in isolated sentences. *Lingua*, *119*(7), 989–1008.

Chung, S. (1982). Unbounded dependencies in Chamorro grammar. *Linguistic Inquiry*, *13*, 39–77.

Chung, S. (1994). Wh-agreement and 'referentiality' in Chamorro. *Linguistic Inquiry*, *25*, 1–44.

Chung, S. (1998). *The design of agreement: Evidence from Chamorro*. Chicago, IL: University of Chicago Press.

Clemens, L. E., Coon, J., Pedro, P. M., Morgan, A. M., Polinsky, M., Tandet, G., & Wagers, M. (2015). Ergativity and the complexity of extraction: a view from Mayan. *Natural Language & Linguistic Theory*, *33*, 417–67.

Clifton, C., Frazier, L., & Deevy, P. (1999). Feature manipulation in sentence comprehension. *Rivista di Linguistica*, *11*(1), 11–39.

Cohen, L., & Mehler, J. (1996). Click monitoring revisited: An on-line study of sentence comprehension. *Memory & Cognition*, *24*(1), 94–102.

Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of agents: The processing of semantic roles in visual narrative. *Cognitive Psychology*, *67*(3), 73–97.

Collins, C., & Thráinsson, H. (1996). VP-internal structure and object shift in Icelandic. *Linguistic Inquiry*, *27*(3), 391–444.

Cooreman, A., Fox, B. A., & Givón, T. (1984). The discourse definition of ergativity. *Studies in Language*, *8*(1), 1–34.

Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, *14*(4), 597–612.

Crain, S., & Fodor, J. D. (1985). How can grammars help parsers? In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 94–128). Cambridge: Cambridge University Press.

de Guzman, V. P. (1995). Experiencers of psych verbs in Tagalog. In C. Burgess & K. D. D. Gerdts (Eds.), *Grammatical relations: Theoretical approaches to empirical questions* (pp. 45–62).

Dickey, M. W., Choy, J. W. J., & Thompson, C. K. (2007). Real-time comprehension of *wh*-movement in aphasia: Evidence from eye-tracking while listening. *Brain and Language*, *100*(1), 1–22.

Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language*, *81*(4), 882–906.

Dillon, B., Staub, A., Levy, J., & Clifton, C. (2017). Which noun phrases is the verb supposed to agree with?: Object agreement in American English. *Language*, *93*(1), 65–96.

Dita, S., & Roxas, R. E. (2011). Philippine languages online corpora: Status, issues, and prospects. In R. E. O. Roxas, S. Hussain, & K.-S. Choi (Eds.), *Proceedings of the 9th workshop on Asian language resources.* Chiang Mai, Thailand: Asian Federation of Natural Language Processing. Retrieved from `https://www.aclweb.org/anthology/W11-3410.pdf`

Drenhaus, H., Saddy, J. D., & Frisch, S. (2005). Processing negative polarity items: When negation comes through the backdoor. In S. Kepser & M. Reis (Eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives* (pp. 145–64). Berlin: Mouton de Gruyter.

Drummond, A. (2016). *Ibex farm: Internet Based EXperiments (Version 0.3.9)* [Computer program]. Retrieved from `http://spellout.net/ibexfarm/`

Dryer, M. S. (2013). Order of relative clause and noun. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online.* Retrieved from `http://wals.info/chapter/90`

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2019). *Ethnologue: Languages of the World* (22th ed.). Dallas, TX: SIL International. Retrieved from `https://www.ethnologue.com/`

Erlewine, M. Y., Levin, T., & van Urk, C. (2017). Ergativity and Austronesian-type voice systems. In J. Coon, D. Massam, & L. deMena Travis (Eds.), *Oxford handbook of ergativity* (pp. 373–96). Oxford, UK: Oxford University Press.

Erlewine, M. Y., Levin, T., & van Urk, C. (2018). *The typology of nominal licensing in Austronesian voice languages.* Poster presented at the 92nd Annual Meeting of the

Linguistics Society of America. Salt Lake City, Utah.

Foley, S., Pizarro-Guevara, J. S., Sasaki, K., Silva-Robles, F., Toosarvandani, M., & Wagers, M. (2019). *Pronouns over gaps in parsing?: Relative clause processing in Santiago Laxopa Zapotec.* Paper presented at the 93rd Annual Meeting of the Linguistics Society of America. New York, NY.

Foley, S., & Wagers, M. (2017). *Subject gaps are still easiest: Relative clause processing and Georgian split ergativity.* Poster presented at the 30th Annual CUNY Conference on Human Sentence Processing. Cambridge, MA.

Foley, W. (1998). Paper presented at the 3rd LFG Conference. Brisbane, Australia.

Foley, W., & Van Valin, Jr., R. D. (1984). *Functional syntax and universal grammar*. Cambridge, UK: Cambridge University Press.

Frazier, L. (1987). Syntactic processing: evidence from Dutch. *Natural Language & Linguistic Theory*, *5*(4), 519–59.

Frazier, L., & Clifton, C. (1987). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, *4*(2), 93–126.

Frazier, L., & Flores d'Arcais, G. B. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, *28*(3), 331–44.

Frazier, M., & Yoshida, M. (2013). *Argument-structure driven parsing in Tagalog.* Poster presented at 26th Annual CUNY Conference on Human Sentence Processing. Columbia, SC.

Gagliardi, A., & Lidz, J. (2014). Statistical insensitivity in the acquisition of tsez noun classes. *Language*, *90*(1), 58–89.

Galang, R. (1988). The language situation of Filipino Americans. In S. L. McKay & S. C. Wong (Eds.), *Language diversity, problem or resource?* (pp. 229–51). Cambridge, MA: Newbury House Publishers.

Galang, R. (2001). Language of instruction in the Philippines in the 20th century: policies, orientations, and future directions. In C. M. Brainard & E. F. Litton (Eds.), *Journey of one hundred years: Reflections on the centennial of Philippine independence* (pp. 97–117). Santa Monica, CA: Philippine American Women Writers and Artists.

Garcia, R., Dery, J. E., Roeser, J., & Höhle, B. (2018). Word order preferences of Tagalog-speaking adults and children. *First Language*, *38*(6), 617–40.

Garcia, R., Roeser, J., & Höhle, B. (2019). Thematic role assignment in the L1 acquisition of Tagalog: Use of word order and morphosyntactic markers. *Language Acquisition*, *26*(3), 235–61.

Garcia, R., Roeser, J., & Höhle, B. (2020). Children's online use of word order and morphosyntactic markers in Tagalog thematic role assignment: An eye-tracking study. *Journal of Child Language*, *26*, 1–23.

Garnsey, S. M., Tanenhaus, M. K., & Chapman, R. M. (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, *18*(1), 51–60.

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language*, *58*(2), 161–87.

Georgopoulos, C. (1985). Variables in Palauan syntax. *Natural Language & Linguistic Theory*, *3*(1), 59–94.

Gibson, E. (1998). The dependency locality theory: A distance-based theory of linguistic complexity. In A. P. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: The MIT Press.

Gibson, E., Desmet, T., Grodner, D., Watson, D., & Ko, K. (2005). Reading relative clauses in English. *Cognitive Linguistics*, *18*(2), 313–53.

Gibson, E., & Wu, H.-h. I. (2013). Processing chinese relative clauses in context. *Language and Cognitive Processes*, *28*(1–2), 125–55.

Gordon, P. (1983). *The acquisition of syntactic categories: The case of the count/mass distinction* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1411–23.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*(1), 97–114.

Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *32*(6).

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, *13*(5), 425–30.

Gordon, P. C., & Lowder, M. W. (2012). Complex sentence processing: A review of theoretical perspectives on the comprehension of relative clauses. *Language and Linguistics Compass*, *6*(7).

Grodner, D. J., & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science*, *29*, 261–91.

Han, C.-H., Lidz, J., & Musolino, J. (2007). V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry, 38*(1), 1–47.

Hawkins, J. A. (1999). Processing complexity and filler-gap dependencies across grammars. *Language, 75*(2), 244–85.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–135.

Hochberg, J. G. (1986). Children's judgements of transitivity errors. *Journal of Child Language, 13*(2).

Hornstein, N. (2009). *A theory of syntax: Minimal operations and Universal Grammar*. Cambridge: Cambridge University Press.

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition, 90*(1), 3–27.

Hsieh, H. (2016). Prosodic indicators of phrase structure in Tagalog transitive sentences. In H. Nomoto, T. Miyauchi, & A. Shiohara (Eds.), *Proceedings of the 23rd Annual Meeting of Austronesian Formal Linguistics Association.* Canberra: Asia-Pacific Linguistics.

Hsieh, H. (2019). Distinguishing nouns and verbs: A Tagalog case study. *Natural Language and Linguistic Theory, 37*(2), 523–69.

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language, 58*(2), 541–73.

Jäger, L., Chen, Z., Li, Q., Lin, C.-J. C., & Vasishth, S. (2015). The subject-relative advantage in Chinese: Evidence for expectation-based processing. *Journal of Memory and Language, 79–80*, 97–120.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*(2), 228.

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133–56.

Kaufman, D. (2009). Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics, 35*(1), 1–49.

Kaufman, D. (2010). *The morphosyntax of Tagalog clitics: A typologically driven approach* (Unpublished doctoral dissertation). Cornell University, Ithaca, NY.

Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, *8*(1), 63–99.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*(5), 580–602.

King, J., & Kutas, M. (1995). Who did what and when? using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, *7*(3), 376–95.

Kroeger, P. (1993). *Phrase structure and grammatical relations in Tagalog*. Stanford: Center for the Study of Language and Information (CSLI).

Kroeger, P. (1998). Clitics and clause structure in tagalog. In M. L. S. Bautista (Ed.), *Pagtanaw: Essays on language in honor of Teodoro A. Llamzon.* Manila: Linguistic Society of the Philippines.

Kroeger, P. (2010). Mckaughan's analysis of Philippine voice. In L. Billings & N. Goudswaard (Eds.), *Piakandatu ami: Dr. howard p. mckaughan* (pp. 207–12). Manila, Philippines: Linguistic Society of the Philippines and SIL Philippines.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Kwon, N., Gordon, P., Lee, Y., Kluender, R., & Polinsky, M. (2010). Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of prenominal relative clauses in Korean. *Language*, *86*(3), 546–82.

Law, P. (2016). The syntax of Tagalog relative clauses. *Linguistics*, *54*(4), 717–68.

Lee, M.-W. (2004). Another look at the role of empty categories in sentence processing (and grammar). *Journal of Psycholinguistic Research*, *33*(1), 51–73.

Levy, R. (2006). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–77.

Levy, R., Fedorenko, E., & Gibson, E. (2013). The syntactic complexity of russian relative clauses. *Journal of Memory and Language*, *69*(4), 461–95.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*, 375–419.

Lim, C., & Erlewine, M. Y. (2018). *Agent extraction and topicalization in Bikol.* Paper presented at the 25th Annual Meeting of the Austronesian Formal Linguistics Association. Academia Sinica: Taipei, Taiwan.

Lin, Y., & Garnsey, S. M. (2010). Animacy and the resolution of temporary ambiguity in relative clause comprehension in Mandarin. In H. Yamashita, Y. Hirose, &

J. L. Packard (Eds.), *Processing and producing head-final structures* (pp. 241–75). Dordrecht: Springer.

Longenbaugh, N., & Polinsky, M. (2018). Equidistance returns. *The Linguistic Review*. doi: http://doi.org/10.1515/tlr-2018-0002

Lowder, M. W., & Gordon, P. C. (2014). Effects of animacy and noun-phrase relatedness on the processing of complex sentences. *Memory & Cognition, 42*(5), 794–805.

Maclachlan, A., & Nakamura, M. (1997). Case-checking and specificity in Tagalog. *The Linguistic Review, 14*(4), 307–33.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Macmillan, N. A., Rotello, C. M., & Miller, J. (2004). The sampling distributions of Gaussian ROC statistics. *Perception and Psychophysics, 66*(3), 406–21.

MacWhinney, B., Bates, E., & Kleigl, R. (1984). Cue validity and sentence interpretation in english, german, and italian. *Journal of Verbal Learning and Verbal Behavior, 23*(2), 127–50.

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language, 47*(1), 50–68.

Mak, W. M., Vonk, W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language, 54*(1), 466–90.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314–24.

Mauner, G., & Koenig, J.-P. (2000). Linguistic vs. conceptual sources of implicit agents in sentence comprehension. *Journal of Memory and Language, 43*(1), 110–34.

Mauner, G., Tanenhaus, M. K., & Carlson, G. N. (1995). Implicit arguments in sentence processing. *Journal of Memory and Language, 34*(3), 357–82.

McKaughan, H. (1958). *The inflection and syntax of Maranao*. Manila, Philippines: Bureau of Printing.

Mitchell, D. C., Cuetos, F., Corley, M., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research, 24*, 469–88.

Nagaya, N. (2019). *Relativization in Tagalog conversation: A typological perspective.* Paper presented at 13th Conference of the Association for Linguistic Typology. Pavia, Italy.

Nagaya, N., & Hwang, H. K. (2018). Focus and prosody in Tagalog. In S. Riesberg, A. Shiohara, & A. Utsumi (Eds.), *Perspectives on information structure in Austronesian languages* (pp. 375–88). Berlin, Germany: Language Science Press. doi: 10.5281/zenodo.1402557

Nagel, H. N., Shapiro, L. P., & Nawy, R. (1994). Prosody and the processing of filler-gap sentences. *Journal of Psycholinguistic Research*, *23*(6), 473–85.

Norcliffe, E., Harris, A. C., & Jaeger, T. F. (2015). Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Language, Cognition, and Neuroscience*, *30*(9), 1009–32.

O'Grady, W., Lee, M., & Choo, M. (2003). A subject-object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition*, *25*(3), 433–48.

Packard, J. L., Ye, Z., & Zhou, X. (2010). Filler-gap processing in Mandarin relative clauses: Evidence from event-related potentials. In H. Yamashita, Y. Hirose, & J. L. Packard (Eds.), *Processing and producing head-final structures* (pp. 219–40). Dordrecht: Springer.

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and language*, *41*(3), 427–56.

Perlmutter, D. M. (1978). Impersonal passives and the Unaccusative Hypothesis. In *Proceedings of the 38th Annual Meeting of the Berkeley Linguistics Society* (pp. 157–89). Linguistic Society of America.

Philippine Statistics Authority. (2017). *Philippines in figures*. Quezon City, Philippines: Republic of the Philippines, Philippine Statistics Authority. Retrieved from `https://psa.gov.ph/sites/default/files/PIF_2017.pdf`

Pizarro-Guevara, J. S. (2014). *The acquisition of Tagalog relative clauses: A comprehension study* (Unpublished master's thesis). California State University, Long Beach, Long Beach, CA.

Polinsky, M., Gallo, C. G., Graff, P., & Kravtchenko, E. (2012). Subject preference and ergativity. *Lingua*, *122*(3), 267–77.

Potsdam, E. (2009). Austronesian verb-initial languages and wh-question strategies. *Natural Language & Linguistic Theory*, *27*(4), 737–71.

Price, I. K., & Witzel, J. (2017). Sources of relative clause processing difficulty: Evidence from Russian. *Journal of Memory and Language*, *97*, 208–44.

Pylkänen, L. (2001). What applicative heads apply to. *University of Pennsylvania Working Papers in Linguistics*, *7*(1), 18.

Qiao, X., Shen, L., & Forster, K. (2011). Relative clause processing in Mandarin: Evidence from the maze task. *Language and Cognitive Processes*, *27*(4), 611–30.

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Rackowski, A. (2002). *The structure of Tagalog: Specificity, voice, and the distribution of arguments* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Rackowski, A., & Richards, N. (2005). Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry*, *36*(4), 565–99.

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, *57*(1), 1–23.

Richards, N. (2009). The Tagalog copula. In S. Chung, D. Finer, I. Paul, & E. Potsdam (Eds.), *Proceedings of the 16th Annual Meeting of Austronesian Formal Linguistics Association.* Canberra: Asia-Pacific Linguistics.

Richards, N. (2017). Some notes on Tagalog prosody and scrambling. *Glossa*, *2*(1), 1–23. Retrieved from `https://doi.org/10.5334/gjgl.252`

Rohde, D. (2003). *Linger: A flexible platform for language processing experiments, Version 2.94.* Retrieved from `http://tedlab.mit.edu/~dr/Linger/`

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic english grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*(3), 348–79.

Ross, J. R. (1967). *Constraints on variable syntax* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Sabbagh, J. (2011). Adjectival passives & the structure of VP in Tagalog. *Lingua*, *121*(8), 1424–52.

Sabbagh, J. (2014). Word order and prosodic structure constraints in Tagalog. *Syntax*, *17*(1), 40–89.

Sabbagh, J. (2016). Specificity and objecthood in Tagalog. *Journal of Linguistics*, *52*(3), 639–88.

Sauppe, S. (2016). Verbal semantics drives early anticipatory eye movements during the comprehension of verb-initial sentences. *Frontiers in Psychology*, *7*, 95. Retrieved from 10.3389/fpsyg.2016.00095

Schachter, P. (1977). Reference-related and role-related properties of subjects. In *Syntax and semantics: Grammatical relations* (Vol. 8, pp. 279–306). New York, NY: Academic Press.

Schachter, P., & Otanes, F. T. (1983). *Tagalog reference grammar*. Berkeley, CA: University of California Press.

Schlesewsky, M., Fanselow, G., Kliegl, R., & Krems, J. (2000). The subject preference in the processing of locally ambiguous wh-questions in German. In *German sentence processing* (pp. 65–93). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Sekerina, I. A. (2003). Scrambling and processing: Dependencies, complexity, and constraints. In S. Karimi (Ed.), *Word order and scrambling* (pp. 301–24). Malden, MA: Blackwell Publishing.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*(1), 103–30.

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, *116*(1), 71–86.

Stowe, L. A. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, *1*(3), 227–45.

Sung, Y.-T., Tu, J.-Y., Cha, J.-H., & Wu, M.-D. (2016). Processing preference toward object-extracted relative clauses in Mandarin Chinese by l1 and l2 speakers: An eye-tracking study. *Frontiers in Psychology*, *7*. doi: 10.3389/fpsyg.2016.00004

Sussman, R. S., & Sedivy, J. (2003). The time-course of processing syntactic dependencies: Evidence from eye movements. *Language and Cognitive Processes*, *18*(2), 143–63.

Tamaoka, K., Sakai, H., Kawahara, J.-i., Lim, H., & Koizumi, M. (2005). Priority information used for the processing of Japanese sentences: Thematic roles, case particles or grammatical functions? *Journal of Psycholinguistic Research*, *34*(3), 281–332.

Tanaka, N. (2016). *An asymmetry in the acquisition of Tagalog relative clauses* (Unpublished doctoral dissertation). University of Hawai'i at Manoa, Honolulu, HI.

Tanaka, N., O'Grady, W., Deen, K., & Bondoc, I. P. (2019). An asymmetry in the acquisition of relative clauses: Evidence from Tagalog. *First Language*, *39*(6), 618–32.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language, 47*(1), 69–90.

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language, 35*(3), 454–75.

Traxler, M. J., Williams, R. S., Blozis, S. A., & Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language, 53*(2), 204–24.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*(3), 285–318.

Ueno, M., & Garnsey, S. M. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processess, 23*(5), 646–88.

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language, 49*(3), 285–316.

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language, 55*(2), 157–66.

Vasishth, S., Brüssow, S., Lewis, R. L., & Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science, 32*(4), 685–712.

Vasishth, S., Chen, Z., Li, Q., & Guo, G. (2013). Processing Chinese relative clauses: Evidence for the subject-relative advantage. *PLoS ONE, 8*(10). doi: 10.1371/journal .pone.0077006

Vikner, S. (2005). Object shift. In H. van Riemsdijk & M. Everaert (Eds.), *The Blackwell Companion to Syntax*. Oxford: Blackwell.

Vincenzi, M. D. (1991). *Syntactic parsing strategies in Italian: The Minimal Chain Principle*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Wagers, M. W., Borja, M., & Chung, S. (2015). The real-time comprehension of wh-dependencies in a Wh-Agreement language. *Language, 91*(1), 109–44.

Wagers, M. W., Borja, M., & Chung, S. (2018). Grammatical licensing and relative clause parsing in a flexible word-order language. *Cognition, 178*, 207–21.

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language, 61*(2), 206–37.

Wagers, M. W., & Phillips, C. (2009). Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics*, *45*(2), 395–433.

Wu, F., Kaiser, E., & Andersen, E. (2012). Animacy effects in chinese relative clause processing. *Language and Cognitive Processes*, *27*(10), 1489–524.

Wu, H.-h. I. (2016). The syntax of correlatives in Isbukun Bunun. *Canadian Journal of Linguistics*, *61*(2), 190–210.

Xiang, M., Dillon, B., & Phillips, C. (2006). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*, *108*(1), 40–55.